Does the Bureaucracy Move the Needle on Trust? Evidence from a Survey Experiment.

Louis Fucilla, Ph.D. Candidate
School of Public and Environmental Affairs & Department of Political Science
Indiana University-Bloomington

*Job Market Paper*

Abstract: This is a stream of discourse in public administration on how best to shape administrative institutions in order to promote trust in government. The underlying assumption is that attitudes towards the bureaucracy are an antecedent of trust and thus by changing attitudes towards the bureaucracy via administration reform one can “move the needle” on trust. Unfortunately, there is limited evidence to support this assumption at the Federal level in the United States. In this paper, I ask several important research questions: does information about Federal agency performance affect trust, is there an asymmetric affect for negative information, and do agency salience or partisan alignment with the presidential administration affect the performance-trust relationship? I address these questions using evidence from a survey experiment that contains three experimental manipulations which vary a cue about agency performance, agency salience and partisan alignment with the presidential administration. The results of the survey experiment suggest that a cue about low agency performance decreases trust but that a cue about high performance does not increase trust relative to a control group exposed to no treatment. I find that there are no interactive effects between the performance cue and agency salience or partisan alignment.
The decline in measures of trust in the Federal government in the United States since the 1960’s is a phenomena of interest to both scholars and practitioners alike. This interest includes efforts to explain the antecedents of trust as well as identify remedies to staunch the decline. Unfortunately, these efforts are not well-integrated. Scholars, mostly in political science, interested in the antecedents of trust tend to look towards attitudes towards elected officials and how citizens’ feel about large policy domains like the economy. Those interested in remedies, often in public administration, tend to focus on changing administration institutions in order to increase trust. While some of these reforms are broad and structural in nature, others have involved increasing use of performance reporting and making information more accessible to citizens. This creates a problem where proposed solutions are not aligned with the known antecedents of political trust. The purpose of this research is to better integrate our understanding of these antecedents of trust in government and the role that public administration might play as a solution. In essence, before we should engage in costly efforts to change administrative institutions as a way to improve trust we should be certain that the bureaucracy does indeed move the needle on trust.

The methodological approach adopted in this analysis is a survey experiment using a factorial design with three treatments and a separate control group. The main research question is whether a cue about the high (or low) performance of one Federal agency increases (or decreases) trust in government. In addition, the experiment also examines the effects of agency salience and partisan alignment with the presidential administration on trust. The results of the experiment suggest that a cue about high agency performance increases trust relative to a cue about low agency performance. Only those receiving a cue about low agency performance had a statistically significant difference in trust than the control group however. Further, the findings
suggest that there were no main and interactive effects for agency salience and partisan
alignment with the performance cue.

The first section of this analysis provides a brief overview of the literature to develop
hypotheses tested in the experiment. This review will address why trust in government is
important, the gaps in the empirical research on trust that this study addresses and how this study
expands upon the growing body of experimental work in public administration on the effects of
performance information and citizen satisfaction with government. The second section provides
a discussion of the survey experiment and its administration via Amazon’s Mechanical Turk or
Mturk. The third section then turns to an analysis of the results and concludes with a discussion
the implications of these findings for both theory about political trust and the feasibility of
administrative reform as a solution to enhance it.

**Trust in Government and Why It Matters**

A full review of the literature on trust and trust in government is beyond the scope of this
analysis. The pertinent issue is what is trust and why does trust in government matter? Defining
trust turns out to be no easy feat and different conceptualizations exist. One common theme is
that trust exists within relationships and thus the common tripartite formulation that “A trusts B
to do, with respect to, X” (Hardin 2006, 19). This three part relationship involves characteristics
of both the truster (A) and the trusted (B) as well as the context of the situation (X) in which trust
is to take place. A truster takes a risk in trusting and leaves themselves vulnerable to the actions
of the trusted (Luhmann 1979, 24). Luhmann argues that this risk taking is important because it
reduces complexity in social interactions and enables cooperation and coordination (1979, 24-
25). In terms of trust in government, this risk taking behavior can be linked to a variety of
important outcomes. In Easton’s conceptualization of diffuse political support, trust is important
because it enables policymakers to govern by creating “… a reservoir of favorable attitudes or
good will that helps members to accept or tolerate outputs to which they are opposed or the
effect of which they see as damaging to their wants” (1965, 273). Extending this logic, more
recent work has focused on the importance of trust in government in securing voluntary
compliance (Braithwaite 1998; Scholz & Lubell 1998; Tyler 1998; Levi & Stoker 2000). For
public management, this suggests that trust in government is important because it eases the cost
of program implementation and administration by reducing the need for more coercion forms of
control and oversight to detect noncompliance.

There is a debate as to whether trust in government is trust in the sense that it exists in
interpersonal relationships. Luhmann distinguishes between interpersonal and system trust which
includes trust in institutions but notes that the latter still functions to reduce complexity as in
interpersonal trust (1979, 48-58). Other scholars adopt the perspective that trust exists in both
“thick” and “thin” forms (Braithwaite 1998; Hardin 2006). Interpersonal trust based on intimate
relationships is an example of “thick” trust. Hardin argues this type of relationship is
unlikely to arise between an individual and government institutions that they little interact with
or know the motivations of the people therein (Hardin 1998, 16) The question then turns to what
exactly is this “thin” form of trust and this is where authors diverge over its importance. For
Hardin, this “thin” form of trust deflates to confidence which is related to but not the same as
“thick” trust (2006, 65). Braithwaite argues that both forms of trust are mutually reinforcing and
lead to voluntary compliance on the part of citizens (1998, 344-345). In another approach, Tyler
argues that “thick” trust may be based on rational calculations of self-interest, but also shared
While this debate is fascinating and conceptually important, this discussion does appear to undercut the logic of the analysis being carried out here. The central question motivating this research is whether information about agency performance affects the commonly used measure of trust in government. This is important because most attempts to increase trust, or confidence, are done via reforms of the bureaucracy and not towards other institutions of government. Even Hardin, a skeptic of trust in government as trust, writes, “Hence, if confidence in government is declining, we should address that problem by enhancing government’s competence, if possible” (2006, 65, also see pages 70-71). This analysis provides a test to see if a cue about the performance of a government agency, i.e. some information that bears on perceptions about its competence, affects trust or confidence in government. Thus one need not privilege the concept of trust in government as it is used here as trust as it exists in interpersonal relationships to accept the utility of this analysis.

**Administrative Reform & Micro-Performance Theory**
The central theoretical mechanism in this analysis comes from Van de Walle and Bouckaert’s micro-performance theory (2003). Van de Walle and Bouckaert write that:

> “Concern with low levels of trust in government and the negative image of government and the public administration has stimulated Western governments to engage in a modernization strategy for their public service. The implicit hypothesis on which this strategy is built is that **better performing public services will lead to increased satisfaction among their users, and this, in turn, will lead to more trust in government.**”

(2003, 892; italics in original).

This logic is consistent with Hardin’s suggestion, mentioned previously, that the solution to declining confidence is to improve government competence. They suggest a model, *Figure 1*, in which agency performance is related to overall trust in government. As agency performance increases so too will satisfaction and trust with that agency and government in general (Van de Walle & Bouckaert 2003, 894). They note here are two key assumptions required for this rational...
to hold: an administrative agency (or agencies) must be associated with government and causality must run from attitudes towards the agency (or agencies) to government in general (Van de Walle & Bouckaert 2003, 895-896). In other words, an agency (or set thereof) is relevant during the process of attitude formation and that attitudes towards that agency (or agencies) must be impactful in influencing evaluations of trust in government. If this relationship holds, then it provides a powerful justification for engaging in administrative reform as a strategy to enhance political trust.

[Figure 1 here]

The evidence to support micro-performance theory at the U.S. Federal level is problematic. There are two studies, Morgeson, VanAmburg and Mithas (2011) and Morgeson and Petrescu (2011), which use attitudes towards specific Federal agencies to explain general trust in government. Both studies use responses to the American Customer Satisfaction Index’s (ACSI) survey. The ACSI includes a sample of Americans who had, or can recall, an experience with a Federal government agency. Items in the ACSI include a measure of general trust in government and indicators of satisfaction and confidence in the specific agencies. Both studies find that customer satisfaction with and confidence in a specific agency affect trust in government (Morgeson, VanAmburg & Mithas 2011, 270; Morgeson & Petrescu 2011, 469).  

1 Van de Walle and Bouckaert do identify a third assumption that the link between an agency’s (or agencies’) performance and trust in government “be direct and linear” (2003, 895). I believe this assumption is not necessary for the internal logic of micro-performance theory and is in fact the hypothesis at the heart of this analysis. As they go on to note, “The main question therefore is what impact the performance of public services has on the evaluations of government in general as compared to other factors” (Wan de Walle & Bouckaert 2003, 895).

2 Van de Walle, Kampen and Bouckaert (2005) provide some evidence using data from Belgium that evaluations of civil servants in general and particular actors and institutions, like politicians, the courts, post officers etc, do affect general evaluations of trust in government. There is a richer body of literature about how micro-performance theory may affect attitudes towards lower-levels of government especially those offered by local governments (see for example: Van Ryzin, Muzzio, Immerwahr, Gulick & Martinez 2004). This analysis focuses only on the U.S. Federal level.

3 Morgeson and Petrescu (2011) conduct both a pooled and agency-specific analysis for the IRS, Medicaid, Medicare, National Park Service, Social Security Administration and Veteran’s Affairs. Their pooled results are similar to the ones found in Morgeson, VanAmburg and Mithas (2011). The results for individual agencies show a
The problem with these studies are related to the ASCI as a data set. The survey does not capture variables for competing explanations of trust in government identified as important in the political science literature including attitudes towards political actors and institutions like the president and Congress, evaluations of highly salient policy areas like the economy and crime and partisanship or political ideology to explain general trust. Additionally, they survey only includes those that have a direct interaction with a Federal agency. While some individuals do interact with the Federal government in this manner, it is more likely that most individuals do so through interaction that is mediated by a third party. This is due to the increasing use of indirect policy tools to provide public goods and services (Mettler 2012). Consequently, a more robust test of micro-performance theory would better mirror real-world conditions in which an individual interacts with or learns about agency performance.

There is a growing body of experimental research in public administration focused on the effects of agency performance using the micro-performance theory paradigm. The experimental approach is fruitful in this case to overcome limitations of a survey like the ACSI. Random assignment to groups ensures that resulting differences in levels of trust are due to the treatment and not other confounding factors (Shadish, Cook and Campbell 2002). These studies tend to focus on local levels of government in both the United States and in other countries but the little more variability. In terms of the relationship between customer satisfaction with the agency and trust, the pooled results appear are driven by experiences with one specific agency. In this case, customer satisfaction with the IRS was related to trust in government, but the relationship did not reach statistical significance in the other five agencies (Morgeson & Petrescu 2011, 469). In terms the relationship between agency confidence and trust in government, the results found in the pooled analysis hold for all but the VA (Morgeson & Petrescu 2011, 469).

In later work, Morgeson does account for the influence of partisanship and political ideology on general trust in government which then functions to influence expectations about Federal agency performance (2013, 293-295). However, this is a significant re-structuring of the logic model of micro-performance theory with causality running from general trust in government to satisfaction and confidence with specific agencies. Such a re-structuring undermines the logic of administrative reform as a means of improving trust in government because trust conditions expectations about agency performance rather than responds to it.
approaches and insights are useful for developing approaches to analyze trust at other levels of government.

Many of these studies focus on examine how citizens’ expectations of government services shape perceptions of and satisfaction with service performance (Van Ryzin 2004, 2006, 2013; James 2011b). Several studies examine how performance information affects satisfaction with government. In one example, James, using both field and laboratory experiments, finds that providing cues that the local government performed well increased satisfaction while cues that it performed poorly decreased satisfaction (2011a, 413). In later work, James and Moseley also found that there was an asymmetric affect for negative performance cues on citizen satisfaction with local government (James & Moseley 2014, 502-503). In their experiment, a negative performance cue lowered satisfaction but a positive performance cue did not lead to a statistically significant increase in higher levels of satisfaction (James & Moseley 2014).

This analysis extends this experimental approach in several important respects to test micro-performance theory. First, the focus is on the United Stated Federal government. Second, it examines how cues about agency performance affect trust in government. Most of the research on trust in government is done at the national level thus this analysis provides results that can speak to that body of literature. Additionally, citizens are more directly connected to local government and less so with the Federal government. This provides a key test of whether performance cues effect trust when the object of evaluation is more remote from direct experience. Further, the cue is provided in a format, a brief news-like narrative, that better mirrors real-world conditions in which most individuals are likely to be exposed to information about Federal agency performance. From the literature, two key hypotheses emerge. The first, Hypothesis 1, is that a performance cue will affect trust. A cue about high performance will
increase trust (H1A) while a cue about low performance will decrease trust (H1B). The second hypothesis (H2) is that a low performance cue will have a larger effect on trust than a high performance cue.

Apart from performance, Van de Walle and Bouckaert’s (2003) analysis of micro-performance theory suggests that the public must connect that agency to the government in order for a performance cue to affect trust. While conceptually straightforward this is an interesting challenge to operationalize in an experimental design. It seems logical that for each individual agencies exist on some continuum of awareness. Some agencies are highly visible with others exist beyond that point where the individual connects that they are indeed a part of the government or cares about that agency. While there may be some shared collective idea of agencies that are more important than others, this continuum is likely to be slightly different for each individual. One approach to operationalizing this continuum would be to adopt the concept of agency salience as a proxy for this continuum of visibility. While several different conceptualizations of salience exist in the political science literature (Wlezien 2005, 557), for the purposes of this analysis salience will be defined as an individual’s judgment about importance the importance of a particular agency. Highly salient agencies are ones that an individual is more likely to connect with government and less salient agencies are less likely to be so connected. Adopting salience as a proxy for visibility suggests that the link between a performance cue and trust might be moderated by the salience of the agency for which the cue is provided (H3). Thus a high performance cue for a highly salient agencies will increase trust more than a similar cue for a less salient agencies (H3A). Inversely, a cue about poor performance for a highly salient agency will decrease trust more than a similar cue for less salient agencies (H3B).
Finally, researchers have also looked at the impact of partisanship on trust. Keele (2005), using American National Election (ANES) survey data and the natural changes in party control after elections, finds that partisans increase trust when their party captures the presidency or takes over a chamber in the Congress. Gershenson, Ladewig and Plame (2006) find a similar affect for partisans after the change in party control of the Senate in 2002. In the experimental literature on agency performance and satisfaction with government, James finds that partisanship partially moderates the relationship between a high performance cue and satisfaction with government (2011a, 414). Hetherington and Rudolph (2014) document how political trust is increasing polarized in the mass public within the last 15 years. Most important to this analysis is Hetherington and Rudolph’s finding that this polarization is partially the result of partisans engaging in motivated reasoning such that they “… increasingly base government evaluations (such as political trust) on the performance criteria that are most (least) favorable to their preferred (non-preferred) party” (2014, 44). In essence, we might expect very complex interactions between performance, partisanship and trust based on motivated reasoning in the process of forming political attitudes.

Motivated reasoning suggests that individuals are motivated to process information in a way that is consistent with established beliefs, values or preferences (Kunda 1990; Lodge & Tabor 2000; Redlawsk 2002). Consider the case of a person who is being given information about the performance of a Federal agency when they share the partisanship of president or aligned with it: if this cue is about good performance, regardless of the salience of the agency, the person might be tempted to increase trust because it confirms existing beliefs or feelings. If the cue is about poor performance, then the person might be tempted to ignore or minimize this information when updating attitudes towards trust. Conversely, consider the situation in which a
person is being given a performance cue when they are not aligned with the partisanship of the presidential administration. In this non-aligned state, the individual may be tempted to minimize the cue about high performance because it is inconsistent with their preferences. This individual may act in the opposite way to a negative performance cue because it is consistent with their preferences. This suggests the hypothesis that subjects will be motivated to adjust their attitudes in a way that is consistent with their partisan preferences (H4). In an aligned condition, subjects will be more likely to increase trust when exposed to a good performance cue than those in a non-aligned condition (H4A). Conversely, subjects in an aligned state are less likely to decrease trust when they receive a poor performance cue that those in a non-aligned state (H4B).

The preceding discussion has focused on developing four testable hypotheses from a large and diverse set of literature about how performance cues, agency salience and partisan alignment may affect trust. A summary of these hypotheses and expectations are in Table 5. The next section discusses the survey experiment conducted to test these hypotheses. While the hypotheses specifically discussed above focus on the main effects of the performance cue and two-way interactive effects for performance and salience and performance and alignment, the analysis will explore the possibility of all main and higher order effects.

Research Design
This study employs a survey experiment using Amazon’s Mechanical Turk (Mturk) workers as subjects. Mturk is a platform that connects a requester, a person with a task they would like performed that requires human intelligence to complete, with a worker willing to perform that task often for a small amount of money. Subjects for this experiment were offered $1 for completing the survey. Because of this structure Mturk is fast becoming a tool used by
social scientists to conduct behavioral experiments (Paolacci, Chandler & Ipeirotis 2010; Berinsky, Huber & Lenz 2012; Goodman, Cryder & Cheema 2013; Paolacci & Chandler 2014). The upside of using Mturk to conduct experiments is that is expands the subject pool beyond those usually available for academic research, like students, and can be run for minimal cost which enables a larger subject pool and increases experimental power (Paolacci, Chandler & Ipeirotis 2010; Berinsky, Huber & Lenz 2012; Paolacci & Chandler 2014). However, Mturk is not a national representative sample and thus the researcher must bear that in mind when assessing the conclusions they can draw from their analysis (Paolacci & Chandler 2014).

Berinsky, Huber and Lenz find that a typical Mturk sample is relatively younger, more Democratic and liberal than the typical nationally representative sample (2012). However, they show that Mturk can be used to replicate previous experimental findings and that “…if we treat the MTurk as a means for conducting internally valid experiments, instead of a representative sample, the MTurk respondent pool is very attractive” (Berinsky, Huber & Lenz 2012, 361).

A survey experiment is “…is nothing more than a deliberate manipulation of the form or placement of items in a survey instrument, for purposes of inferring how public opinion works in the real world” (Gaines, Kuklinski & Quirk 2007, 3). This experiment involves a 2x2x2 factorial design with an independent control group where the treatment factors being manipulated are the cue about performance (high or low), the salience of the agency contained in the performance cue (high or low) and the alignment with the presidential administration mentioned in the performance cue (aligned or not). Overall, there are 8 treatment groups plus the control group. The outcome variable is the standard ANES trust in government index.

While survey experiments are widely used in political science, one limitation of this analysis does not consider the duration of the treatments on trust. Gaines, Kuklinski and Quirk
(2007) argue that duration is important if we wish to understand public opinion in the real world. However, due to cost limitations, it was only possible to measure the outcome variable immediately after treatment and thus this is the immediate effect of the treatments on trust. Additionally, the cue was written in a manner that performance is unambiguously high or low. Information about agency performance may well be more ambiguous in cues provided by in the real world thus this provides evidence about the likely maximal effect of a cue on trust (Gaines, Kuklinski & Quirk 2007). However, the experiment does employ an independent control group that receives no treatment which allows for a comparison between treatment and control groups which is identified as a best practice in survey experiment research (Gaines, Kuklinski & Quirk 2007).

An overview of the experimental protocol is provided in the following discussion and the full instrument including performance narratives is available in Appendix I. A potential subject enters the survey via the Mturk platform and after the Institutional Review Board (IRB) study page is screened with the standard party identification and leaning partisan questions. To minimize cost, this survey only includes subjects who identified as partisans or reported leaning towards one party or the other. There is also good evidence that there are relatively few “true” independents in the population (see for example Petrocik 2009). Once subjects were screened, they were randomly assigned to the treatment or control group. All subjects then received a set of items asking about their political ideology and support for President Obama, the Congress and Federal departments and agencies in general. The order of these items was randomized as were response options when possible. After these items, subjects in the control group received a narrative about the effects of winter weather during 2013-2014 and then the ANES trust in
government index followed by standard socio-demographic items like age, gender and educational attainment.

Subjects assigned to an experimental condition were asked to complete an activity prior to receiving the experimental treatment. These subjects were provided information about the mission of six Federal agencies. These agencies were the Environmental Protection Agency (EPA), Social Security Administration (SSA), Consumer Product Safety Commission (CPSC), Department of Transportation (DOT), Department of Commerce (Commerce) and the Federal Communications Commission (FCC). Information about each agency’s mission was gathered from the agency’s website and given to subjects verbatim. Subjects in treatment groups where then asked to rank the importance of each agency’s mission to them personally or the agency’s valence. Subjects were then asked to rank order the agencies on a scale of 1 to 6 in terms of how important they thought the agency’s mission is to the “well-being or good of the nation”. This ranking served as the measure as the subject’s perception of agency salience or importance. While valence and salience are related concepts, the decision was made to use “well-being or good of the nation” question as the measure of agency salience since research in other areas of political behavior suggest the importance of sociotropic attitudes versus individual considerations (see for example Kinder & Kiewiet 1981).

Subjects in the treatment conditions were then assigned a narrative that manipulated the following information: agency performance (high or low), agency salience (high or low based on the subject’s own ranking) and partisanship of the presidential administration so that it was

---

5 These agencies were selected because they are relatively well-known and reflect the natural distribution of agency preference estimates from Clinton and Lewis’s work (2008). These estimates are comparable to common factor ideal point estimates popular in the study of Congress (Clinton & Lewis 2008). The distribution of agency preference skews more liberal so the decision was made to include more agencies with more liberal preference estimates. This includes the CPSC (-1.69), EPA (-1.21) and SSA (-0.45). One agency, DOT (0.07), is relatively towards the center in terms of preference estimates while the FCC (0.32) and Commerce (1.25) represent more conservative agency preferences.
aligned (or not) with the partisanship of the subject. Treatment groups are thus comprised of a mix of both Republicans and Democrats (including partisan leaners) who received a narrative with a presidential administration that was consistent (or not) with their treatment condition. The choice to lump Republican and Democrats into the same treatment groups was made for several reasons. First, having separate groups would have required 8 additional treatment groups which was cost prohibitive. Second, given the characteristics of Mturk workers as a subject pool (e.g. more Democratic), it would have been harder to find roughly equal groups of partisans.

The performance narrative consisted of three cues about an agency’s performance in terms of its: ability (or not) to meet its goals or objectives, have high (or low) levels of customer satisfaction as measured by the ASCI and have a reported high (or low) level of employee morale per a report from the Office of Personnel Management (OPM). All of these performance cues were framed as relative in that they compared the performance of the agency mentioned in the narrative to other Federal agencies thus an agency was either one of the best or top performing Federal agencies or one of the worst or lowest performing agencies. James argues that relative performance information provides an easy cue for citizens to grasp that does not require complex information processing (2011a, 401). One benefit of this approach is that it provides an inclusive bundle of potential measures of agency performance which can be complex and multi-dimensional (Rainey 2003). It includes a measure of effectiveness in goal attainment as well as the perception of external stakeholders in customer satisfaction. It also includes an internal measure of performance in the cue about employee morale.

---

6 Republicans and Democrats were separately assigned to groups to ensure that both sets of partisanship had the same probability of being assigned to a group.
7 For instance, in an aligned group all Republicans would receive a narrative with a Republican president and all Democrats would receive a narrative with a Democratic President. This was flipped for non-aligned groups so that Republicans received a narrative with a Democratic President etc.
One issue that arises here that other experimental work avoids is the issue of deception. For this experiment, it was necessary to create narratives that were deceiving as actual performance is more nuanced in the real world. In his work on agency performance, James and colleagues (2011a&b, w/Moseley 2014) specifically focus on conducting experiments in low and high performing areas where citizens are exposed to only information specific to their particular local context. This avoids the issue of deception because subjects are only provided information about good (or bad) performance if it is indeed how their local government has performed. There are varying norms in different social science fields about deception and one of the concerns is that the use of the deception “poisons” a potential subject pool and can hurt the validity of the study (Sell 2008; Barrera & Simpson 2012). Using an experiment to test the negative claims about validity, Barrera and Simpson find that deception does not harm the validity of the study (2012). Subjects in the treatment condition were provided information at the end of the survey that notified them of the deception so that they were not left with the impression that the performance cue they received was actual fact.

After receiving the performance narrative, subjects in the treatment condition received an attention check question. The question was designed to see if subjects had paid attention to the narrative and could correctly identify whether it was about good (or bad) performance. Roughly 94% of subjects in the treatment groups passed this question. Subjects in both treatment and control groups where then asked to answer the standard items contained in the ANES trust in government index which is the outcome variable of this study and a memory dump question designed to probe what subjects were thinking about when they answered the question. Subjects in the treatment group had two additional questions about the performance narratives and their
agency salience rankings. Lastly, subjects had a small battery of socio-demographic questions and then the survey ended.

Overall there were 3,045 subjects. 304 subjects comprise the Control group with the remaining subjects divided among the 8 Treatment groups. Table 2 below provides some descriptive statistics about participants in the final subject pool. Additionally, it shows the results of testing whether participants in the treatment and control groups differed based on party identification, political ideology, attitudes towards political actors and institutions as well as socio-demographic characteristics. The results suggest that groups do not vary from one another at conventional levels of statistical significance (α=0.05) on most characteristics.

[Table 2 here]

Initially the subject pool was supposed to be 2,700 subjects or roughly 300 subjects per group. The survey experiment initially ran late in late May of 2015. After collection and cleaning of the data, it was discovered that there was an error in the programming of the survey that lead to issues with two Low Performance, Low Salience treatment group. The Low Performance, Low Salience and Aligned Partisan group was only composed of Republicans while the Low Performance, Low Salience and Non-Aligned group was only composed of Democrats. The programming error lead subjects to being randomly being assigned to 7 instead of 8 treatment groups. This lead to slightly higher numbers of subjects per group (approximately 340 per group except for the control). The decision was made to re-run the four low performance treatment groups and this completed in late June of 2015. While there is the possibility that time may have affected the results, the new Low Performance groups are very similar to the original groups they are replacing. Table 3 highlights the statistically significant differences between the original and replacement groups. Two of the groups have no differences and two have differences related to
evaluations of the Federal bureaucracy in general and towards Congress. Perhaps most importantly none of the original or replacement groups varied in terms of mean level of trust. This should provide some confidence that the replacements groups are indeed comparable replacements.8

[Table 3 here]

**Findings from the Survey Experiment**

Responses to the survey experiment are analyzed using ANOVA and the regression tables produced by that procedure. The results for the three-way full factorial model comparing only treatment groups are displayed in Model 1 in *Table 4*. The results suggest that there is no three-way interaction between the treatment conditions.9 Additionally, none of the two-way interactions are statistically significant. The Wald F-test statistic (1, 2733) testing for the equality of the marginal means for the two-way interaction between the Performance and Salience treatments is 0.01 with a p-value of 0.92. This indicates that I cannot reject the null hypothesis that the means are equal to one another. The Wald F-test statistic (1, 2733) testing for the two-way interaction between the Performance and Alignment treatments is 0.21 with a p-value of 0.65. Again, I cannot reject the null hypothesis that the means for these groups are equal to one another. Overall this indicates that I find no support for the Hypothesis 3 or 4 about the interactive effects between the Performance and Alignment treatment conditions.

---

8 Additionally, the findings presented in the next section are consistent when the analysis is conducted and different choices about including subjects and groups from the first and second runs of the experiment are made. For instance, the results are the same if the two Low Performance groups unaffected by the programming error (High Salience and Aligned Partisanship and High Salience and Unaligned Partisanship) in the first run are included and those the same groups in the second run are excluded. Additionally, the results are also the same if you include only those subjects from the second run that serve to fill the gaps as a result of the programming error (i.e. including only Democrats for the Low Performance, Low Salience and Aligned group and only Republicans for the Low Performance, Low Salience and Unaligned group).

9 Beyond the coefficient produced by the ANOVA regression, I test for the hypothesis that the marginal means of the 8 treatment groups are equal. The Wald F-test statistic (1, 2733) is 0.00 with a p-value of 0.95. This indicates that I cannot reject the null hypothesis that marginal means for all three way interactions are equal.
Indeed the only hypothesis to receive any support is for Hypothesis 1 about the main effects of performance. The results provide support for both H1A, that a good performance cue increases trust, and H1B, that a poor performance cue decreases trust. Form the result in Model 1, receiving a high performance cue increases trust by slightly less than 4 points on the 100 points trust in government index. I test this by conducting a Wald F-test that the cell means are equivalent. The test statistic (1, 2733) is 5.53 with a p-value of 0.02. This indicates I can reject the null hypothesis that the means between the low and high performance cue groups are equal. Additionally, I estimate the effect size for the performance cue, partial $\omega^2$, which is interpretable as the amount of variance explained in trust by this variable. Performance does have a small, but non-zero effect on trust. Another way of approaching this is to look at graphs for the marginal means for the two way interactions which are in *Figure 2*. The results show that trust increases when subjects are assigned a high performance cue and this occurs regardless of the salience or alignment treatment condition. Regardless of whether the agency is less salient or the partisanship of the presidential administration is not aligned with their own, subjects exposed to a high performance cue increase trust relative to those in a low performance treatment cue.

The question now turns to whether these results hold when comparing the performance cue groups to the control group which received no treatment and whether there is an asymmetric affect for a negative performance cue (H2). To do this I ran another ANOVA using just the performance treatment and the results are displayed in *Table 5*. The results from this model suggest that the subjects in treatment groups who receive a high performance cue are no different
than control subjects which offers disconfirming evidence for H1A.\textsuperscript{10} However, the results for the model offer support H1B and H2. A negative performance cue does decrease trust and this decrease in trust gives subjects in these groups a mean level of trust that is different from trust in the control\textsuperscript{11} and high performance cue groups\textsuperscript{12}.

[Table 5 here]

Lastly, I provide a check of robustness of the findings by examining subjects the top of the head considerations after responding to the Trust in Government index. These were collected as qualitative comments from subjects.\textsuperscript{13} If the treatments did affect trust, then subjects assigned to the treatment groups should be more likely to mention the bureaucracy in general or a specific agency. I coded comments so that 0 means that the bureaucracy, either in general or a specific agency, was not mentioned while 1 indicates that it was. Table 6 displays the number of subjects falling into these categories after coding for the control group and all treatment groups. The Pearson $\chi^2$ test of association suggests that there is an association between whether subjects mention the bureaucracy or not in the responses and being in the control group or one of the treatment groups. Additionally, I estimated a probit model with this Mention variable as the dependent variable and Group (0 if control and 1 if in a treatment group) as the independent variable.\textsuperscript{14} I then computed the change in the predicted probability of mentioning the

\textsuperscript{10} More formally I test the null hypothesis that the coefficients for the control and high performance treatment groups are equal. The Wald F-statistic (1, 3042) is 0.03 with a p-value of 0.86. This indicates I cannot reject the null hypothesis that the means for the groups are equal.
\textsuperscript{11} I test the null hypothesis that the coefficients for the control and low performance cue treatment groups are equal. The Wald F-statistic (1, 3042) is 6.27 with a p-value of 0.01.
\textsuperscript{12} I test the null hypothesis that the coefficients for the low and high performance cue treatment groups are equal. The Wald F-statistic (1, 3042) is 14.92 with a p-value in excess of 0.00.
\textsuperscript{13} The text of the item from the survey was: “Thinking about the last few questions you just answered, please list a few people, organizations or ideas that came to mind as you were answering those questions. You do not need to provide a detailed answer or in-depth response. Just a few words or a short responses is sufficient.”
\textsuperscript{14} Results of the model are not shown. The Likelihood Ratio $\chi^2$ for the model was 6.03 with 1 degree of freedom and a p-value of 0.014. N was 3,045 and the Pseudo $R^2$ was 0.002. The coefficient for the Group variable was 0.205 with a standard error of 0.084 and a p-value of 0.015.
bureaucracy based on assignment to the control group versus one of the treatment groups. The results indicated that subjects in a treatment group were approximately 27%^15 more likely than those in the control group to mention the bureaucracy in general or a specific agency as a top of the head consideration when responding to the Trust Index.

Overall, the findings provide evidence that a negative performance cue does decrease trust relative to a control group and a high performance cue group. Further because the high performance cue group and the control group have similar mean leaves of trust this suggests that there is an asymmetric effect for negative performance information. None of the hypothesis concerning the interactive effects between performance and salience or alignment treatment conditions (H3 and H4) were supported in the analysis. Finally, the robustness check of comments made after the Trust Index suggests that respondents in the treatment groups were influenced, in part, by considerations about the bureaucracy when answering that battery of questions. I now turn to a discussion of the findings.

**Discussion**

The overarching research question addressed in this study was whether the performance of Federal agencies affects trust in government. The results are novel in that they are the first to provide experimental evidence that this is indeed the case. At first blush this would seem to suggest that reforming administrative institutions to make them more competent might be a viable strategy to enhance trust. However that interpretation would be misleading. Positive performance information does not increase trust and subjects receiving a cue about good performance were no different than a control group.

15 The 95% CI for the estimate of the predicted probability is 0.259 – 0.291.
This does suggest that there is indeed a potential to actually decrease trust via administrative reform. Proponents of administrative reform are eternally optimistic about the efficacy of reform. Unfortunately, the history of administrative reform in the public-sector is usually one of mixed results which Beryl Radin attributes to the grafting of private-sector management practices onto public agencies existing in very different environmental contexts (2012, 36-37). In such instances, failure to live up to expectations may cause trust to decrease.

These findings also speak to some expectations about the effect of agency performance in the public administration literature. In a critical review of micro-performance theory by Yang and Holzer (2006). They argue that there exists a “…tenuous relationship between individual agencies' performance and CTG\textsuperscript{16} at the national level” (Yang & Holzer 2006, 116). However, they suspect that jurisdiction-wide performance, how government is doing across a wide variety of concerns, is likely to strongly affect trust affect (Yang & Holzer 2006, 117). The results suggest that individual Federal agencies have a non-negligible effect on trust in government which disconfirms their expectation. While the effect is small, these results show that individual agencies do effect trust in government.

However, within the subject pool, one of the biggest predictors of trust in government is how the subject generally feels about attitudes towards the Federal agencies. Figure 3 shows the interactive effects from an ANOVA of performance treatments and responses to this question at the beginning of the survey instrument. It is important to remember that all treatment groups do not vary significant between each other. In essence, subjects that have really positive attitudes towards Federal agencies were equally likely to be assigned to any group. The results show that subjects with more positive attitudes towards Federal agencies in general have high levels of

\textsuperscript{16} This is their acronym for citizen trust in government.
trust regardless of the performance cue they receive. Subjects with more negative attitudes towards Federal agencies in general are less trusting regardless of the performance cue they receive. Thus this provides some evidence to confirm Yang and Holzer’s expectation that these government-wide assessment are significant to the development of trust in government.

[Figure 3 here]

This suggests several avenues for future experiment work. First, researchers should see if cueing information about Federal agencies in general (e.g. for example by cueing information that agencies generally perform well or not) affects trust in government. Additional experimental evidence that there is a link between general attitudes towards Federal agencies and trust would provide evidence to support this conclusion. Second, researchers may wish to test whether cueing agency performance information affects general attitudes towards Federal agencies. While the effect of any one agency on general trust in government may be small, provision of a low or high performance cue may exert a large effect on attitudes towards agencies in general which may then affect trust.

The failure of the experiment to generate significant findings for the interactive effects is disappointing. This could be due to two reasons: low power or experiment design. In general, we can rule out statistical power as an explanatory variable. I conducted a power analysis prior to the experiment to determine the number of subjects needed for a given level of power. Because there was little prior research on this topic, I estimated the smallest detectable mean difference between groups as 0.32 (on a 100-point index). With a group size of 300, the power to detect the 3-way interaction was in excess of 0.99. Power only increased when the size of the treatment groups was increased to approximately 340. Thus the experiment is sufficiently powered to detect even minute differences between groups.
That leaves the other explanation, experimental design, as the likely culprit. In this case, I think there are issues with the salience and alignment treatments. In terms of salience, my assumption would be that rank ordering six agencies would produce enough variance to create an interactive effect with trust. A likely problem is that subjects did not have strong enough feelings that the lowest ranked agency was significantly less salient than their highest ranked agency. If this is the case, the design can likely be revised to better capture this notion of salience in order to provide a better test. In terms of alignment, I suspect that the cue was not sufficient to trigger a partisan response. In terms of the current political environment, a better cue might have been to say agency X under President Obama or agency X under a Republican Congress rather than simply asking about an agency under a Democratic or Republican presidential administration. However, the results do show the utility of including a control group that receives no treatment. Had the control group not been included then I would have made misleading inferences about the effects of a high performance cue.

Finally, there are a few limitations that need to be noted. First, this analysis provides only a look at the effect of a performance cue on trust immediately after the cue has been received. As noted by Gaines and colleagues (2007) and experiment should also examine the duration of the effect. Thus this study cannot say that this decrease in trust from those receiving a performance cue lasted beyond the end of the experiment itself. Second, the experiment provides a likely maximal immediate effect given the straightforward nature of the performance cue. Future research might adopt approaches in the framing literate in political science to explore how mixing competing, in this case, cues affect the outcome variable (see for instance Chong & Druckman 2007). Third there was an error in the administration of the survey initially that lead to issues with assignment to the treatment groups. The logic of the survey experiment is that
subjects receive treatment at roughly the same time without other factors intervening. This is not the case, but I have endeavored to show that the decision to re-run the low performance groups did not appear to create systematic differences between the groups particularly with respect to the outcome variable.

Overall my results show that a cue about agency performance does affect trust and this makes a contribution to the literatures on trust in government, experiments in public affairs and citizen responses to agency performance. The results suggest that a negative cue about performance produces a small, but significant decrease in trust relative to a control that that received no cue and even a group that received a high performance cue. This provides evidence that there is an asymmetric effect for negative information. Overall attitudes towards government, their antecedents and the impact they have on the implementation and administration of public policy is a key area where the fields of public management and political science overlap. Findings, such as those presented in this analysis, suggest that the bureaucracy need to be incorporated into theory and research on trust in political science. This is lacking in the existing literature and is an important omission. This analysis also underscores the applicability of using survey experiments, more common in political science, to answer questions relevant to key concerns in the study of public management. There is a body of scholarship that focuses on bureaucratic politics and thus speaks to both fields with its findings. However this is primarily based on the study of the bureaucracy as an institution. This analysis presents another way of drawing connections between these fields while primarily drawing from theories about political behavior. It represents another way of approaching the study of bureaucracy that is distinct from but complements the institutional approach. This approach is novel and there exists
room for more inquiry and blending of methodological approaches across the fields of public management and political science.
References


Appendix 1: Survey Instrument

INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR
Attitudes towards the U.S. Federal Government

You are invited to participate in a research study of attitudes towards the United States Federal Government. You were selected as a possible subject because you are a Worker who responded to the study’s HIT posted on Amazon’s Mechanical Turk website. I ask that you read this form and ask any questions you may have before agreeing to see if you are eligible to take part in the study and then participating in it. The study is being conducted by Louis Fucilla, a graduate student of the School of Public and Environmental Affairs and the Department of Political Science at Indiana University-Bloomington, with oversight by his advisor, Professor Sergio Fernandez.

STUDY PURPOSE
The purpose of this study is to examine attitudes towards the Federal Government in the United States.

PROCEDURES FOR THE STUDY:
If you agree to be in the study, you will do the following things: You must be eligible to participate in the study. To check whether you are eligible to take the survey, you will be asked two quick screening questions at the beginning of the survey which will take less than a minute to answer. Those eligible will be directed to the survey. Those who are not eligible will not receive a payment.

If you are eligible to participate in the study, you will complete a short survey which includes questions about attitudes towards elected officials and other parts of the U.S. Federal government, read a short article and then answer some follow-up questions about attitudes towards government and some demographic questions. The total time to participate in the study should not exceed 15 minutes.

CONFIDENTIALITY
If you are eligible to participate in the study and choose to do so, your name will not be linked to your survey responses at any time. I do not collect your name in the survey and therefore not even the researcher will know who provided which sets of responses.

Organizations that may inspect and/or copy your responses for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review Board or its designees, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP).

PAYMENT
You will receive payment a $1.00 payment for taking part in this study if you are eligible to participate in and complete the study. Eligibility will be determined by two quick screening questions at the beginning of the study as mentioned above and you will not receive payment if you are not eligible to participate in the study. Additionally, you need to complete the study. This means that you need to complete the required questions in order to receive a confirmation code to enter into the HIT to receive payment. If you do not wish to complete the required questions, then I am not able to accept your response and you will not receive payment for participating in the study.
CONTACTS FOR QUESTIONS OR PROBLEMS
For questions about the study, contact the researcher, Louis Fucilla, at (812)822.1872 or lfucilla@indiana.edu.

For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at (812) 856-4242 or (800) 696-2949.

VOLUNTARY NATURE OF STUDY
Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with the School of Public and Environmental Affairs, the Department of Political Science or Indiana University-Bloomington.

Check on the box below if you would like to print a copy of this sheet for your own records.

Eligibility Screening Questions
Please answer the following questions which will determine your eligibility to participate in the study.

Generally speaking, do you usually think of yourself as a:
- Republican
- Independent or Other
- Democrat

If Republican/Democrat Selected
Would you call yourself a strong (D/R) or a not very strong (D/R)?
- Strong
- Not Very Strong

Answer If Generally speaking, do you usually think of yourself as a: Independent or Other Is Selected
Do you lean more to the Republican party or to the Democratic party?
- Republican
- Independent or Don't lean towards either party
- Democratic
- Don't know

End of Survey Message if Ineligible to Participate

If Eligible
You are eligible to participate in the study based on your responses to the last 2 questions. Please be sure to read the questions carefully and provide responses when you are required to do so.

When it comes to politics, do you usually think of yourself as:
- Extremely Liberal
- Liberal
- Slightly Liberal
- Moderate or Middle of the Road
- Slightly Conservative
- Conservative
- Extremely Conservative
Do you approve or disapprove of the way Barack Obama is handling his job as President?
- Strongly Approve
- Approve
- Neither Approve or Disapprove
- Disapprove
- Strongly Disapprove
- Don't Know

Do you approve or disapprove of the way the U.S. Congress has been handling its job?
- Strongly Approve
- Approve
- Neither Approve or Disapprove
- Disapprove
- Strongly Disapprove
- Don't Know

How much trust and confidence do you have in the agencies and departments of the federal government when it comes to carrying out the functions of the federal government:
- A great deal
- A fair amount
- Not very much
- None at all

Additional items assigned only to Subjects in a Treatment Condition.
The next few questions focus on some departments and agencies of the United States Federal government. These organizations, along with their missions, are listed below.

Environmental Protection Agency (EPA): to protect human health and the environment.

Social Security Administration (SSA): deliver Social Security services that meet the changing needs of the public.

Federal Communications Commission (FCC): to make available, so far as possible, to all the people of the United States, without discrimination on the basis of race, color, religion, national origin, or sex, rapid, efficient, nationwide, and worldwide wire and radio communication service with adequate facilities at reasonable charges.

Department of Commerce: to create the conditions for economic growth and opportunity.

Department of Transportation (DOT): to serve the United States by ensuring a fast, safe, efficient, accessible and convenient transportation system that meets our vital national interests and enhances the quality of life of the American people, today and into the future.

Consumer Product Safety Commission (CPSC): to protect the public against unreasonable risks of injury from consumer products through education, safety standards activities, regulation, and enforcement.
How important are the mission statements for the departments and agencies listed above to you personally? I mean, how much do you personally care about the mission of each department or agency?

- Environmental Protection Agency (EPA)-Ranked on a scale from 1.Not Important at All to 7. Extremely Important
- Social Security Administration (SSA)—same as above
- Federal Communications Commission (FCC)—same as above
- Department of Commerce—save as above
- Department of Transportation—same as above
- Consumer Product Safety Commission—same as above

Again, thinking about these departments and agencies of the U.S. Federal government and their missions. Please rank the departments and agencies in terms of how important you think their mission is to the well-being or good of the nation on a scale from 1 to 6 where 1 is Most Important and 6 is Least Important.

While the missions of some of these agencies may be more or less important to you than others, you may only rank one department/agency as Most Important and one as Least Important.

The list below is randomly generated. To rank the departments and agencies, drag and drop the boxes with their names so that the Most Important one is ranked "1" and the second is ranked "2" until you have the Least Important one ranked as "6".

_____ Environmental Protection Agency (1)
_____ Social Security Administration (2)
_____ Federal Communications Commission (3)
_____ Department of Commerce (4)
_____ Department of Transportation (5)
_____ Consumer Product Safety Commission (6)
## Control and Treatment Narratives

<table>
<thead>
<tr>
<th>Control</th>
<th>High Performance</th>
<th>Low Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Winter of 2013-2014 will not be remembered kindly but most Americans. Residents of the vast swathes of the country experienced prolonged periods of bitter cold and record amounts of snowfall. The weather closed schools and business and affected travelers during the holidays. As most of the nation huddled together for warmth, areas of the West and Southwest experienced one of the warmest winters on record. While they missed out on the cold, residents in these areas are now dealing with a crippling drought brought on by lack of precipitation during those unusually warm winter months. This drought has, in turn, affected the rest of the country by raising the price of agricultural goods from those states. In short, let’s hope for a milder winter next year in the Plains, Midwest and East and a wetter one in the West and Southwest.</td>
<td>News about the X (Most/Least Salient Agency)</td>
<td>News about the Federal government in Washington, D.C. can often be bad, but not for the X under a Republican (Democratic) Administration. According to a recently released report, the X met or exceeded its performance goals in the last year. This made it one of the highest performing agencies, top 10%, in the Federal government. Additionally, the annual American Customer Satisfaction Index (ACSI), which tracks customer satisfaction with the Federal government, identified X as one of the Federal agencies who have the most satisfied customers. Last, in a survey of employees conducted by the Office of Personnel Management (OPM), X was identified as one of the best places to work in the Federal government. X’s employees felt that their co-workers very dedicated to their jobs and worked hard for the good of the American public.</td>
</tr>
</tbody>
</table>

### Attention Check question for Treatment Subjects
Thinking about the story you just read, would you say that the agency's performance was:

- Good
- Poor
- Don’t Remember

### Outcome Variables & Open-ended Memory Dump Question All Subjects Receive
People have different ideas about the government in Washington. These ideas don’t refer to Democrats or Republicans in particular, but just to government in general. We want to see how you feel about these ideas. For example:

How much of the time do you think you can trust the government in Washington to do what is right:

- Just about always
- Most of the time
- Only some of the time
- Almost never
Would you say the government is pretty much run by a few big interests looking out for themselves or that it is run for the benefit of all the people?
- A few big interests
- Run for the benefit of all

Do you think that people in the government waste a lot of money we pay in taxes, waste some of it, or don’t waste very much of it?
- Waste a lot of money
- Waste some of it
- Don't waste very much of it

Do you think that quite a few of the people running the government are crooked, not very many are, or do you think hardly any of them are crooked?
- Quite a few people running the government are crooked
- Not very many are crooked
- Hardly any of them are crooked

Thinking about the last few questions you just answered, please list a few people, organizations or ideas that came to mind as you were answering those questions. You do not need to provide a detailed answer or in-depth response. Just a few words or a short responses is sufficient.

Extra Items for Treatment Group Subjects
In general, if you were going to evaluate the performance of a Federal department or agency how important are each of the following?
- Attaining goals—Ranked on a Scale from 1.Completely Unimportant to 10.Extremely Important
- Customer satisfaction—same as above
- Employee satisfaction and commitment—same as above

Earlier in the survey you ranked the X (Most Salient Agency) as the most important Federal agency among the agencies that were listed. Thinking about your rankings, what are a few ideas that came to mind as you were ranking these agencies? Your responses do not need to be long or in-depth and a few words or phrases are sufficient. The complete list of agencies you were asked to rank is provided below: (All Agencies Displayed)

End of Survey Socio-Demographic Questions for all Subjects
You are close to finishing the survey. Please take a minute or two and provide some basic demographic information about yourself:

How old are you?
- 18-25
- 26-34
- 35-54
- 55-64
- 65 or over

What is your gender?
- Male
- Female
What is the highest level of education you have completed?
- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

What is your annual income range?
- Below $20,000
- $20,000 - $29,999
- $30,000 - $39,999
- $40,000 - $49,999
- $50,000 - $59,999
- $60,000 - $69,999
- $70,000 - $79,999
- $80,000 - $89,999
- $90,000 or more

**Deception Reveal for Treatment Subjects**
You are almost done!

Before you are finished, please be aware that the information you were provided in the story about the performance of an agency was created by the researcher and does not reflect that agency’s actual performance in the last year.

Reports from the American Customer Satisfaction Survey (ASCI) and the Office of Personnel Management (OPM) do exist and can be accessed online if you are interested in seeing how these groups assess the performance of Federal departments and agencies.

The purpose of this study is to assess whether getting information about a well-performing agency (or not), getting this information about an agency that was more (or less) salient to you and for a presidential administration that was aligned (or not) with your own partisanship affects general trust in the Federal government.

Feel free to contact the researcher responsible for this study, Louis Fucilla, at (812)822-1872 or lfcilla@indiana.edu if you have any questions or concerns.

Lastly, please refrain from discussing the particulars of this study with other workers in the Mechanical Turk community until after the study closes. This research is part of Louis’s dissertation and he needs the experiment to run smoothly so he can collect his data and finish his dissertation.

**End of Survey Message w/Payment Code**
Tables and Figures

Figure 1: A Logic Model for Micro-Performance Theory

![Logic Model Diagram]

Note: Reproduction of Figure 1 in Van de Walle and Bouckaert (2003, 894)

Table 1. Hypotheses Developed from Literature Review

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Statement of Expected Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Main Effect of Performance</td>
</tr>
<tr>
<td></td>
<td>A: High Performance Cue ↑ Trust</td>
</tr>
<tr>
<td></td>
<td>B: Low Performance Cue ↓ Trust</td>
</tr>
<tr>
<td>2</td>
<td>Asymmetric Effect of Low Performance</td>
</tr>
<tr>
<td></td>
<td>(Low Performance Cue ↓ Trust &gt; High Performance Cue ↑ Trust)</td>
</tr>
<tr>
<td>3</td>
<td>Interactive Effect between Performance and Salience</td>
</tr>
<tr>
<td></td>
<td>A: High Performance Cue for High Salience Agency ↑ Trust &gt; High Performance Cue for Low Salience Agency ↑ Trust</td>
</tr>
<tr>
<td></td>
<td>B: Low Performance Cue for High Salience Agency ↓ Trust &gt; Low Performance Cue for Low Salience Agency ↓ Trust</td>
</tr>
<tr>
<td>4</td>
<td>Interactive Effect between Performance and Partisan Alignment</td>
</tr>
<tr>
<td></td>
<td>A: Aligned and High Performance Cue ↑ Trust &gt; Non-aligned and High Performance Cue ↑ Trust</td>
</tr>
<tr>
<td></td>
<td>B: Aligned and Low Performance Cue ↓ Trust &lt; Non-aligned and Low Performance Cue ↓ Trust</td>
</tr>
</tbody>
</table>
Table 2. Descriptive Statistics of Final Subject Pool

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Significant Differences between Groups (α=0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partisanship (1=Republican, 2=Democrat)</td>
<td>1.70</td>
<td>0.46</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Political Ideology (7-Category)</td>
<td>4.76</td>
<td>1.72</td>
<td>1</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>Presidential Approval (0=Disapprove/Don’t Know, 1=Approve)</td>
<td>0.48</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Congressional Approval (0=Disapprove/Don’t Know, 1=Approve)</td>
<td>0.11</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>General Trust &amp; Confidence in Federal Agencies (4-Category)</td>
<td>2.42</td>
<td>0.65</td>
<td>1</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>Age (5-Category)</td>
<td>2.29</td>
<td>0.97</td>
<td>1</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>Education (8-Category)</td>
<td>4.28</td>
<td>1.37</td>
<td>1</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>Sex (0=Male, 1=Female)</td>
<td>0.52</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>Yes²</td>
</tr>
<tr>
<td>Income (9-Category)</td>
<td>3.67</td>
<td>2.49</td>
<td>1</td>
<td>9</td>
<td>-</td>
</tr>
</tbody>
</table>

Note 1: To test for differences between groups I employed chi squared tests of association for the above variables and a variable for each Group (Control and Treatment). I also employed logistic and ordinal logistic regression using these variables as the dependent variable and the subject’s assignment to a particular treatment and control group variable as the independent variable.

Note 2: The results for a logistic regression suggests Treatment Group#3 (High Performance, High Salience, Non-Aligned) is significantly more male than female relative to the control group. The chi squared test of association finds that there is no association between Sex and Treatment group assignment. In general, a subject’s sex does not predict variation in trust.

Table 3. Significant Differences between Original & Replacement Groups (α=0.05)

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Difference</th>
<th>Mean Original (Std)</th>
<th>Mean Replacement (Std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Performance, High Salience &amp; Aligned</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Performance, Low Salience &amp; Aligned</td>
<td>General Trust &amp; Confidence in Federal Agencies (4-Category)</td>
<td>2.33 (0.70)</td>
<td>2.09 (0.69)</td>
</tr>
<tr>
<td>Low Performance, High Salience &amp; Non-Aligned</td>
<td>Congressional Approval (0=Disapprove/Don’t Know, 1=Approve)</td>
<td>0.08 (0.27)</td>
<td>0.13 (0.33)</td>
</tr>
<tr>
<td>Low Performance, Low Salience &amp; Non-Aligned</td>
<td>None</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Model 1: Three-way Full Factorial ANOVA for Treatment Groups

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Coefficient (Std. Error)</th>
<th>p-value</th>
<th>( \omega^2 ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>3.72 (1.58)</td>
<td>0.02</td>
<td>0.005 (0.001-0.012)</td>
</tr>
<tr>
<td>Salience</td>
<td>1.52 (1.58)</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Alignment</td>
<td>2.02 (1.57)</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td><strong>Two-Way Interactions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PerformanceXSalience</td>
<td>-0.23 (2.23)</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>PerformanceXAlignment</td>
<td>-1.03 (2.23)</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>SalienceXAlignment</td>
<td>-1.67 (2.22)</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td><strong>Three-Way Interaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PerformanceXSalienceXAlignment</td>
<td>-0.22 (3.14)</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>20.51 (1.11)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td><strong>Model Diagnostics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F (7, 2733)</td>
<td></td>
<td>2.55</td>
<td>0.01</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td>2,741</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Effect of Performance Cues over Other Treatments
Table 5. Model II: Main Effects for Performance for Treatment & Control Groups

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Coefficient (Std. Error)</th>
<th>p-value (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Performance Cue</td>
<td>-3.27 (1.31)</td>
<td>0.01 (0.001-0.011)</td>
</tr>
<tr>
<td>High Performance Cue</td>
<td>-0.23 (1.31)</td>
<td>0.86</td>
</tr>
<tr>
<td>Constant</td>
<td>25.13 (1.18)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Model Diagnostics
- F (7, 2733) = 8.45, p = 0.00
- R² = 0.00
- n = 3,045

Table 6. Mentions of the Bureaucracy in Treatment and Control Groups after Trust Index

<table>
<thead>
<tr>
<th>The Bureaucracy is:</th>
<th>Control n (%)</th>
<th>Treatment n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Mentioned</td>
<td>238 (78%)</td>
<td>1,968 (72%)</td>
</tr>
<tr>
<td>Mentioned</td>
<td>66 (22%)</td>
<td>773 (28%)</td>
</tr>
<tr>
<td>Total</td>
<td>304</td>
<td>2,741</td>
</tr>
</tbody>
</table>

Pearson χ² is 5.77 with 1 degree of freedom and a p-value of 0.02
Figure 3. Interaction between Performance Treatments and General Attitudes about Federal Agencies with 95% CIs