INCREDIBLY GOOD PERFORMANCE: AN EXPERIMENTAL STUDY OF SOURCE AND LEVEL EFFECTS ON THE CREDIBILITY OF GOVERNMENT

ABSTRACT

Initiatives to boost public trust of government often rely on better reporting of the efforts and accomplishments of government agencies. But if citizens disbelieve the performance reports of agencies, especially information about good performance, then these initiatives will be ineffective. We ask: Do citizens find performance information from government agencies to be credible, or do they trust more in independent sources? Do they believe some agencies more than others? And does credibility of the agency itself as a source depend on the level of performance that is being reported? To address these questions, we designed an experiment to test the credibility of a customer satisfaction index for two US federal agencies, with random allocation of the specific agency (one politically less attractive, the other more so), the source of the index (the federal agency itself, or an independent ratings firm), as well as the level of performance reported in the index. Results from an online sample of nearly 600 US adults show that credibility is clearly lower for the politically less attractive agency , and that citizens are especially doubtful about good performance reported by the government agency itself (as opposed to the independent ratings firm). These results suggest that independent sources can boost credibility when reporting good news about government performance.

INTRODUCTION

Many citizens think that government agencies do not do a good job in terms of achieving the objectives set for them (Hibbing and Theiss-Morse, 2001). Low performance contributes to lack of trust in government with citizens having perceptions, attitudes and behaviors that discourage them from putting what they value, including their time and money, in the hands of government (Gambetta, 1988; Rousseau et al., 1998). Distrust on the part of citizens lowers support for redistributive policy programs (Hetherington, 2005) and inhibits cooperation with government activities and laws (Tyler, 1990). The reporting of performance information to improve public accountability to citizens is a core part of arguments for enhancing citizens' trust of government (Yang & Holzer, 2006). This idea, part of a broader attempt to make government more open and transparent, has been promoted by influential non-governmental organisations, including (in the US) the Government Accounting Standards Board, the International City/County Management Association, and the National Association of State Budget Officers, and (internationally) Transparency International and the National Democratic Institute. Governments around the world have been attempting to increase transparency by providing more information about inputs, processes, outputs and outcomes (Hood, 2006; James, 2011; Grimmelikhuijsen. and Meijer, 2012). To take the US as a prominent example, the transparency requirements of the Government Performance and Results Act were strengthened in 2010 with a requirement that agencies make their performance information publicly available on a website (performance.gov), and the UK Government published performance against a set of Public Service Agreements for government activity in the 2000s.

To date, there has been little research on whether citizens find the reporting of government agencies' performance to be credible. In particular, citizens may be sceptical of information about government agencies' performance of their tasks when it is reported by those agencies themselves. Organisations have an interest in presenting their performance positively; several forms of unreliability and lack of validity, including deliberate misrepresentation, have been

documented (Hood, 2008; 2011; Moynihan, 2008). General distrust of government among the public may compound the lack of credibility (Chanley, Rudolph and Rahn, 2000; Alford, 2001; Dalton, 2005; Van de Walle, Van Roosbroek, and Bouckaert, 2008; Van Ryzin, 2011). There is some evidence suggesting that simple performance measures provided by government organisations can be as credible to citizens as information from other sources (Van Ryzin and Lavena, 2013), but we do not know if this finding applies to more complex services and whether it applies in different jurisdictional contexts. Additionally and especially, information about high levels of government performance may be particularly likely to lack credibility amongst citizens because of concerns about government agencies' incentives to show they are performing well.

In this paper we ask: Do citizens find performance information from government agencies to be credible, or do they trust more in independent sources? Do they believe politically more attractive agencies, serving clients seen as deserving, above politically less attractive agencies, serving clients viewed as less deserving? And does credibility depend on the level of performance, particularly when the government itself reports higher performance relative to lower performance? In the next section we outline theoretical expectations of how the source and level of reported performance may influence credibility. A subsequent section sets out the experiment we designed to evaluate our hypotheses. We then present our findings, which show that credibility is lower for a less popular federal agency, and that citizens are especially doubtful about good performance reported by the government agency itself, as opposed to an independent ratings firm. We conclude by drawing out the implications of these observed credibility effects for government performance reporting and democratic accountability.

CREDIBILITY AND SOURCES OF PERFORMANCE INFORMATION

There is a great deal of published information about the performance of public organisations and public services in many contemporary contexts, reporting inputs, processes, outputs or outcomes (Hatry, 1999; Moynihan, 2008; James, 2011; Van Ryzin, 2011). These efforts to report on government performance aim at enhancing the reputation of government agencies in the

eyes of citizens by making public service efforts and accomplishments more transparent and visible, counteracting the often negative image of government performance in society (Goodsell, 2003). The potential for such attitude change, however, rests on the assumption that citizens will find government performance reports to be credible, which is the basic assumption we test here. We define credibility as a combination of the perceived truthfulness of the level of performance reported on a measure and the perceived believability of the measure itself as a way of gauging performance. Both these elements reflect citizens' concerns with honesty and competence as a basis for trust and reflect the definition of credibility consisting of these elements that has been used extensively in political psychology (Franzoi, 1996, following on from Hovland & Weiss, 1951).

Research on information and change in audiences' beliefs and attitudes suggests that the source of information matters for forming beliefs and underpinning consequent attitudes and behaviour (Pornpitakpan, 2004). The expertise and competence of the source, having the interests of the audience at heart (and not exploiting them), the attractiveness or likeability of the source, and the similarity between the source and the audience have all been identified as factors affecting credibility (Yoon et al., 2011). In the context of information about public service performance, some of these factors are likely to affect whether citizens trust information from government agencies about their performance. In particular, because in many contexts citizens find government performance information somewhat abstract and difficult to judge, we would expect them to rely on heuristics or summary cues when forming their judgments about the information's credibility. Thus, government performance information may engage citizens' quick response, sometimes called 'peripheral', route for perception and attitude change, which is typical for many political contexts where citizens do not have a personal, deep involvement in the issue (Lodge and Taber, 2013). In public services, this processing would be typical of perceptions of services that citizens do not themselves regularly use. This peripheral route contrasts with the central route for perception and attitude change which is more cognitively

taxing and time consuming, involving more systematic, deep reflection and processing of information (Petty and Cacioppo, 1986; Eagly and Chaiken, 1993).

Information in the form of a summary cue about local government performance from an independent auditor has been found to be credible to citizens in previous experiments, changing both their perceptions of performance and satisfaction (James, 2011; James and Moseley, 2014). But to date there has been little research contrasting different sources of government performance information. In an experiment about reporting street cleanliness, Van Ryzin and Lavena (2013) found very little difference in the credibility of a graphical report of the percentage of acceptably clean streets provided by government itself, compared with either a nongovernmental source or a no-named source. These authors suggest that credibility effects may be more likely when the service is less tangible or the measurement is more abstract, so that sources act as a cue when citizens form their trust judgments. We test source effects precisely in this context, where such effects are potentially more likely to be observed. Thus, for our experiment, we have chosen to look at an index of overall customer satisfaction for government agencies, which is a fairly abstract, global performance measure that depends a great deal on how satisfaction is defined and evaluated. Related research has examined the effects of different forms of transparency, defined as release of information, on whether public organisations are seen as competent, acting in pursuit of the interests of affected people, and whether they tell the truth (Grimmelikhuijsen and Meijer, 2013). However, this work did not vary the source of the information. In our study, we look directly at citizens' perceptions of the credibility of performance information itself, and also citizens' assessments of actual performance relative to reported performance, to assess the effect of different sources of information on citizens' judgements.

There are several potential alternatives to a government agency reporting on its own performance, including performance monitoring by independent public or private sector audit bodies, non-profit watchdog organizations, and private rating firms. For our study, we have

chosen to study independent rating firms. These are independent from government agencies in that they are not publicly owned by the government, do not rely mainly on government funding, and are not regulated by government in their measurement and reporting activity (so they are not regulated in the way that some economic activity is subject to rules run by state regulatory bodies). There are many rating agencies but our focus is on a major private firm that has developed a measure of user satisfaction with public and private services, the American Customer Satisfaction Index (ACSI). ACSI developed its satisfaction index originally in Sweden (Andersen, Fornell and Lehman, 1994) and later applied the methodology extensively in the US to measure customer satisfaction with a wide range of private and public services, including services provided by several US federal government agencies. Although these government agencies contract with ACSI to have their services included, the ACSI methodology is highly standardized across both the private and public sectors and thus individual firms and agencies cannot influence ACSI's measurement and reporting of their satisfaction ratings.

The resulting satisfaction index information is available publicly and published online both by ACSI and by the various federal government agencies as part of their broader performance reporting efforts. This allows us to frame the information as coming from either ACSI or a federal agency itself. When customer satisfaction is reported by the agency itself, given general distrust in government and concerns specifically about performance measures, we might expect such information to be seen as less credible. This expectation leads to our first, and most basic, hypothesis:

> • H1: Citizens will be less trusting of performance information when government is reporting on itself (compared to a private firm as an independent source).

Even when government is reporting on itself, as prior research on information source effects suggests, the political attractiveness of the agency to the audience may influence reception of the message. Attractiveness is often defined as perceived familiarity, likability, and similarity of the source to the receiver in terms of being part of group identity (McGuire, 1985). In the

6

context of government agencies, more attractive agencies that enjoy broader public support because of favourable attitudes toward their programs or clients may be more likely to be seen as trustworthy sources of information, compared to government agencies that are less attractive. In our study, we focus on two US federal agencies actually covered by the ACSI: the Veterans Administration (VA), the main agency serving veterans (a politically popular client group serving predominantly US citizens); and the Citizenship and Immigration Services (CIS), the main agency serving immigrants (a less politically popular client group who are predominantly non-US citizens). Although we recognize that these agencies differ in many complex ways, and we do not have an empirical gauge of attractiveness, we choose the VA and CIS from among the federal agencies included in ACSI as contrasting agencies in terms of their assumed attractiveness in the eyes of the American public because of their very different client groups. Thus, we suggest the following hypotheses about the attractiveness of the source and credibility:

- H2: Citizens will be less trusting of performance information from a politically less attractive agency (in this study, CIS) compared to a politically more attractive agency (in this study, VA).

Not all performance reporting by an agency on itself paints the same picture: it can reveal better or worse news about agency efforts and accomplishments, and we would expect that the perceived credibility of the information will vary in response. The literature on performance reporting suggests that citizens are likely to be more sceptical of relatively positive information about performance than information about relatively lower performance. Government sources have an incentive to misrepresent their performance as being high (in order to gain rewards or avoid criticism), but there is less obvious reason to think that they would misrepresent their performance as being low. Thus, we hypothesize that this suspicion of misrepresentation is likely to be most salient when the agency reports good news about itself. Citizens' scepticism about an agency self-reporting good performance has some foundation in reality, it should be

noted, as there are many examples of cases where performance has been misrepresented by government agencies, including exaggerating or even falsifying the information in order to look good (Bevan and Hood, 2006).

Thus, we expect that when agencies report positive information about their own performance it will be seen as less credible by citizens. In turn, when agencies report less positive information about their own performance it is more likely to be believed. By varying the positivity of the information about performance of the agency (by showing higher and lower performance), and by varying the source (government itself versus a more independent source), we can thus test the following interaction hypothesis:

- H3: When the level of reported performance is higher citizens will be less trusting of government reporting on itself relative to when reported performance is lower. In other words, the interaction between government as a source and a higher level of reported performance will have a negative effect on credibility.

EXPERIMENTAL DESIGN AND PARTICIPANTS

The experiment we designed to test these hypotheses presents customer satisfaction information about a federal government agency with random allocation of three elements: the source of the information, the agency, and the level of the satisfaction index. Appendix A provides the complete text and images used in an online survey to present the source, agency and index information.

We modelled the presentation of the information after ACSI, which as discussed above is a private firm that independently tracks customer satisfaction, with reported satisfaction index scores theoretically ranging from 0 to 100 (with 100 representing complete satisfaction). Both VA and CIS contract with ACSI to track their customer satisfaction, and in recent years they have

received similar ratings (in the high 80s to low 90s). As mentioned previously, the agencies and ACSI separately report this information to the public on their websites, enabling us to realistically frame the index as coming from either the government agency itself or from ACSI, which was described to respondents as an independent organization (since ACSI is not widely known and recognized). This manipulation of source (agency versus ACSI) allowed us to test the central question of whether an agency reporting on itself is seen as less credible (Hypothesis 1).

The random assignment of two different agencies enabled us to evaluate Hypothesis 2, namely, that the VA (which serves veterans and is thus more politically attractive) would be seen as more credible than the CIS (which serves immigrants and is less politically attractive). Each agency's logo was included on the page and the full name of the agency was used, rather than the acronym, in the short written presentation of the satisfaction index (again see Appendix A). It should be noted that this study was conducted several months before the most serious and highest profile revelations of mismanagement and misconduct at the VA emerged in May, 2014, eventually resulting in the resignation of the VA Secretary Shinseki. Still, the emerging scandal surrounding the VA at the time of our study may have attenuated the difference in its political attractiveness, as compared with CIS. However, more broadly, it is important to note that whilst using two real agencies is desirable because it makes the study relevant, the two bodies are complex bundles of activities and citizen perceptions of them, so this is a partial test of the issue of perceived agency characteristics.

To evaluate the source by level of performance interaction hypothesis (Hypothesis 3), we randomly varied the presented value of the 0-100 satisfaction index within a range of whole numbers from 86 to 93, based on the real range of ACSI scores reported for our two agencies. This allowed us to examine whether credibility declined with higher reported performance, compared to lower performance, especially when the agency was reporting on itself. We contrasted scores in the range 86-89, defined as "lower" (left digit 8), with scores in the range 90-93, defined as "higher" (left digit 9). Although all these satisfaction scores are based on the

real ACSI scores for both agencies and fairly high overall, in absolute terms, the difference still allows us to exploit the left digit bias effect. Left digit bias has been found in a wide variety of contexts, including perceptions of government performance information (Olsen 2013), and accentuates the lower-higher contrast because recipients of the information concentrate on the left digit of the numbers.

For all three randomly varied factors (source, agency, and level), we tested for covariate balance using measures of age, sex, race, education, income, political ideology, and trust of government (see Appendix B). No significant differences were found across factors at the .10 level using the Bonferroni correction, thus indicating successful randomization on these variables.

There are two dependent variables that come from a series of questions that directly followed the presentation of the randomized source, agency and satisfaction index. The first is a 'credibility scale' composed of two items: a question asking, *Do you think the level of customer satisfaction reported for the agency is . . .* with responses from *1=definitely not true, to 5=definitely true*; and a question asking *In your view, how believable is this index as a way of measuring the satisfaction of actual customers with the agency's services?,* with responses from *1=not believable at all, to 10=completely believable.* The combination of these two questions into a scale reflects our conceptualization of credibility as involving both honesty and competence. For ease of interpretation, the credibility scale was transformed to a 0-100 range (mean = 47.4, SD = 25.3), and the scale has good reliability (alpha = .86). The second dependent variable is a single question that asked: *What do you think the true level of the satisfaction index really is -- your best guess?* Respondents could indicate a response from *1=much worse, to 7=much better* than the reported satisfaction index (again, a random value from 86 to 93), which was inserted at the midpoint of the 1-7 response set. This question thus captures respondents' judgment of the extent to which the satisfaction index under- or over-states the agency's real customer satisfaction. To facilitate interpretation, we transformed this measure to a -3 to 3 scale (mean =

-.86, SD = 1.33), with negative values indicating a belief that the real customer satisfaction is lower than the reported level and positive values indicating a belief that it is higher.

Participants were adult residents from across the US recruited using CivicPanel, a university-affiliated online research panel, and the survey was conducted in March, 2014. A total of 8,832 panelists were sent an email invitation, of which 1680 were confirmed to have opened the email invitation and 689 responded after two contact attempts (representing a 19 percent contact rate and a 41 percent cooperation rate). We dropped 49 non-US respondents as well as 41 respondents who completed the survey in under 2 minutes (because of concerns that they sped through the questions and did not fully read and consider the information presented to them), leaving an analytical sample of 599. Partial completion and item nonresponse resulted in somewhat smaller samples in the analyses, depending on the variables involved. Participants were predominately non-Hispanic white (81%) and disproportionately female (67%), with an average of 49 years of age, 15 years of education, and $75,000 of annual income. In terms of political ideology, they are fairly balanced (mean = 5.5 on a 1-10 left-right scale), and they have levels of trust of government that are slightly below the midpoint of the 0-100 scale and comparable to national statistics from the American National Election Studies (Gershtenson & Plane, 2007). Appendix B provides the precise statistics for the sample overall as well as across experimental factors.

Our experiment involved randomly varyingthe ACSI scores shown to respondents. The range of these scores was narrow and quite high overall (86-93) and, as mentioned, was based on the real range of ACSI scores for both agencies. The random variation used meant that not all the particular scores given to individual participants were exactly as reported in ACSI. However, the variation from the actual ACSI scores was small and harmless (in terms of potentially misleading participants about the true performance of these agencies), particularly in the light of the ACSI score itself likely having some small measurement error as a measure of the true satisfaction of the customers of these agencies. The experimental variation of the source was

consistent with actual practice, as the ACSI scores are in fact reported by both the agencies themselves and the ACSI organization.

RESULTS

Figure 1 presents descriptive results for our 0-100 credibility index in the form of a graph of means by treatment group status. Overall, the perceived credibility is somewhat low, with all but one of the means falling below the midpoint of the 0-100 scale. Only the difference between agencies, CIS versus VA, is marginally significant (p < .10, two-tailed test), with CIS being judged to be somewhat less credible than VA. Although the difference is not significant statistically, the reported customer satisfaction information appears slightly less credible when framed as coming from the agency itself as opposed to the independent source ACSI. Figure 2 presents the means by treatment group for the question that asked respondents about their best guess as to the true level of customer satisfaction. As can be seen, in general respondents thought the true level was lower, implying that the satisfaction index information presented to them overstated the true level of customer satisfaction. Although the mean for the less attractive CIS appears somewhat lower overall, neither the source nor the agency effects are significant statistically (even at the .10 level).

To formally test the main effects, as well as to examine interactions, Table 1 presents the results of a regression analysis including the agency factor (CIS or VA), the source factor (the agency itself or ACSI), and the randomized level of the satisfaction index. To make the OLS regression coefficients equivalent to main effects, the agency factor and source factor were both effect coded as follows: agency (-.5=VA, .5=CIS) and source (-.5=ACSI, .5=the government agency itself). The satisfaction index level was dichotomized, as mentioned earlier, to capture left-digit bias and also effect coded (-.5=86-89, .5=90-93). Table 1 presents results for the two dependent variables, showing models both without and with the two-way interaction terms. Although substantively we are interested primarily in the source by level (left digit) interaction, the other

two possible interactions (agency x source and agency x left digit) were also included in the model to be complete and to avoid bias. (The three-way interaction terms were not of substantive interest and not statistically significant in either model, so they were not included.)

Model 1 in Table 1 shows just the main effects on credibility. Paralleling the descriptive results from Figure 1, the source effect is negative as hypothesized (Hypothesis 1), but the effect is small and not statistically significant. Only the agency (CIS) effect is marginally significant ($p <$ .10, two-tailed test), with CIS being viewed as somewhat less credible than the VA, as expected from our attractiveness hypothesis (Hypothesis 2). There is no evidence of a left digit main effect. Model 2 adds the three two-way interactions, one of which is highly significant statistically: when the agency is reporting on itself—and the satisfaction index level is higher (left digit = 9)—respondents judged the information to be much less credible. This is the interaction of relevance to our interaction or positivity hypothesis (Hypothesis 3), and the results confirm our expectation that citizens seem especially sceptical of government-provided information about good performance.

Model 3 in Table 1 shows just the main effects for our second dependent variable, regarding the respondents' best guess about the true satisfaction level. Although the source and agency effects are in the hypothesized direction, neither is statistically significant, as mentioned earlier when presenting Figure 2. The left digit main effect is again also not significant. Model 4 includes the interactions, and the results indicate that the interaction of source by left digit is negative and statistically significant ($p < .05$, two-tailed test). Consistent with our interaction or positivity hypothesis (Hypothesis 3), when an agency is reporting a higher satisfaction index for itself, respondents tend to believe that the real level of customer satisfaction is lower, implying that the agency is over-stating its performance.

Figure 3 helps illustrate our finding of an interaction of source (the agency reporting on itself) and a higher level of reported satisfaction (left digit = 9), which again is the interaction of interest for our positivity hypothesis (Hypothesis 3). The figure is based on Model 2 in Table 1

13

and refers to the credibility scale. Interestingly, when the reported satisfaction index is relatively lower (left digit = 8), there is little difference in the credibility of the reported information, and indeed the credibility appears slightly higher even when the agency is reporting on itself. But when the reported satisfaction score is relatively higher (left digit = 9), there is a distinct drop in the credibility of the information when the agency itself is the source. The drop in credibility is about 9 points on the 0-100 credibility scale, which is fairly substantial given an overall mean level of about 47. The strength of this interaction effect is especially noteworthy given that difference between "lower" and "higher" customer satisfaction was only the left digit of a reported index with a relatively narrow range (from 86 to 93). We discuss the implications of these findings in the next section.

DISCUSSION AND IMPLICATIONS

Our experimental results provide at best only a weak indication of a source main effect, with respondents viewing the reported customer satisfaction as slightly less credible when the agency was reporting on itself, as compared with a more independent source (the ACSI), although the differences were not statistically significant. This lack of much difference between government versus non-government sources is consistent with the one previous experiment along similar lines (Van Ryzin and Lavena, 2013). Still, we designed our experiment to probe this previous finding more deeply by focusing on the federal government and using the American Customer Satisfaction Index (ACSI), a reputable and independent firm, and thus expected to see more of a source main effect. One explanation for the lack of a large source main effect may be that, despite being a leader in its field, ACSI is perhaps not well known enough to have much of an independent reputation in the minds of most citizens (such as the well-known consumer rating firm JD Power, for example).

Despite finding only a weak main effect, however, our results show a very large interaction effect of source by level of reported performance. This suggests a different, more complex

interpretation of the way in which citizens judge the credibility of government-reported performance. When the agency was reporting on itself—and when the level of reported customer satisfaction was higher (left digit = 9)—participants were much less likely to judge the information as credible, and they were more likely to think that the agency was over-stating its performance. In contrast, when the reported level of customer satisfaction was relatively lower (left digit = 8), agency self-reporting was found to be more credible. This interaction effect was substantively large and has potentially important implications. For one thing, this interaction effect suggests a problem facing an agency that performs well and has good news to report: namely, that citizens may be especially distrustful of government-provided information precisely when an agency most expects to receive credit and earn public trust by reporting high performance. Paradoxically, elected officials and agency executives often strive to report their performance in highly positive terms, sometimes bending the results to present their efforts and accomplishments in the most positive light possible (Hood, 2006). Indeed, it may well be the suspicion of such framing or gaming tactics that breeds distrust of glowing government performance information in the first place. Our results suggest that such practices, rather than helping government, may only erode the trust and credibility of the government performance movement more generally in the eyes of citizens.

An important implication of our interaction finding between the agency as a source and the level of performance is to call attention to the value of an outside or independent rating of government performance, especially when performance is high. A development of this line of research would be to consider alternative forms of involvement by bodies independent of the agency whose performance is being reported, for example to consider a hybrid case where the agency presents the information but the information is independently certified or in other ways validated by an independent organisation. In this way, the achievements of government may be made more credible to citizens.

Our experimental results provide partial support for our hypothesis about the attractiveness of a government agency and its influence on credibility, with the VA viewed as somewhat more credible than CIS. This finding is consistent with Heatherington's (2005) research showing that Americans express less trust of government agencies that serve out-groups and require sacrifice. However, the agencies differ in many respects, not only their clients, so more research on other dimensions is merited, for example the extent of redistribution involved in the agency's programs, citizen knowledge about the agency, or the extent of media coverage of the agency. Veterans Affairs experienced a large amount of negative media coverage (relative to historical coverage) in the months around the time our study was conducted. Despite this caveat, the findings suggest that citizens' trust of government performance information likely varies by the prestige and political popularity of the particular agency or arm of government doing the reporting. Less prestigious or less politically attractive agencies face an uphill struggle in their effort to demonstrate their efforts and accomplishments to the public.

*   *   *

REFERENCES

Alford, J. R. (2001). We're all in this together: The decline of trust in government, 1958–1996. In J. R. Hibbing & E. Theiss-Morse (Eds.), What is it about government that Americans dislike? (pp. 28–46). Cambridge: Cambridge University Press.

Anderson, E. W., Fornell C., & Lehman, D. R. (1994). Customer satisfaction, market share and profitability: Evidence from Sweden. Journal of Marketing, 58(3), 53-66.

Bevan, G. & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English healthcare system. Public Administration, 84(3), 517-538.

Chanley, V. A., Rudolph, T. J., & Rahn, W. M. (2000). The origins and consequences of public trust in government: A time series analysis. Public Opinion Quarterly, 64(3), 239-256.

Christensen, T., & Lægreid, P. (2005). Trust in government: The relative importance of service satisfaction, political factors, and demography. Public Performance & Management Review, 28(4), 487-511.

Dalton, R. J. (2005). The social transformation of trust in government. International Review of Sociology, 15(1), 133-154.

Eagly, A. H. & Chaiken, S. (1993). The Psychology of attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Gershtenson, J., & Plane, D. L. (2007). Trust in government. American National Election Studies Pilot Report. Retrieved from http://www.electionstudies.org/resources/papers /Pilot2006/nes011890.pdf

Gallup (2014). Gallup World Affairs Series, Gallup.

Gambetta, D. (1988). Can we trust? In D. Gambetta (Ed.), Trust: Making and breaking cooperative relations (pp. 213-38). New York, NY: Blackwell.

Goodsell, C. T. (2003). The case for bureaucracy: A public administration polemic (4th ed.). Washington, D.C: CQ Press.

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

Hatry, H. P. (1999) Performance measurement: Getting results. Washington DC: Urban Institute Press.

Hetherington, M. J. (2005). Why trust matters: declining political trust and the demise of American liberalism. Princeton, N.J.; Woodstock: Princeton University Press.

Hibbing, J. R. & Theiss-Morse, E. (2001). Introduction: Studying the American people's attitudes toward government. In J. R. Hibbing, & E. Theiss-Morse (Eds.), What is it about government that Americans dislike? (pp. 1-9). Cambridge: Cambridge University Press,.

Hood, C. (2006). Gaming in targetworld: the targets approach to managing British public services. Public Administration Review, 66(4), 515-521.

Hood, C. (2006). Transparency in historical perspective. In C. Hood and D. Heald (Eds.), Transparency: The key to better governance? (pp.3-23). Oxford: Oxford University Press.

Hood, C. (2011). The blame game: Spin, bureaucracy, and self-preservation in government. Princeton and Oxford: Princeton University Press.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. Public Opinion Quarterly, 15(4), 635-650.

James, O. (2011). Performance measures and democracy: Information effects on citizens in field and laboratory experiments. Journal of Public Administration Research and Theory, 21(3), 399-418.

James, O. & Moseley, A. (2014). Does performance information about public services affect citizens' perceptions, satisfaction and voice behaviour? Field experiments with absolute and relative performance information. Public Administration, 92(2), 493-511.
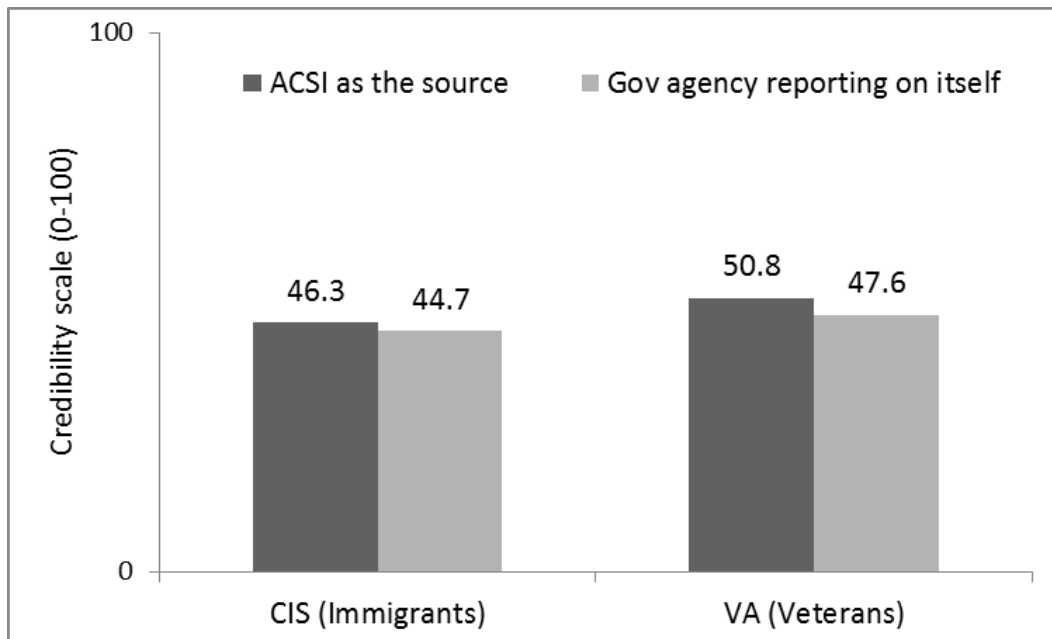
Lodge, M., & Taber, C. S. (2013). The rationalizing voter. Cambridge ; New York: Cambridge University Press.

McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), Handbook of social psychology (Vol. 2, pp. 233-346). New York: Random House.

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

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.

Petty, R. E. & Cacioppo, J. T. (1986). Communication and persuasion: Central and peripheral routes to attitude change. New York: Springer-Verlag.

Pew Research Centre (2013). United Nations retains strong global image. Washington D.C.: Pew Research Centre.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. Journal of Applied Social Psychology, 34(2), 243-281.

Rousseau, D., Sitkin, S., Burt, R., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. Academy of Management Review, 23(3):393-404.

Tyler, T. R. (1990). Why people obey the law. New Haven: Yale University Press.

Van de Walle, S., Van Roosbroek, S., & Bouckaert, G. (2008). Trust in the public sector: Is there any evidence for a long-term decline? International Review of Administrative Sciences, 74(1), 45-62.

Van Ryzin, G. G. (2011). Outcomes, process, and trust of civil servants. Journal of Public Administration Research and Theory, 21(4), 745-760.

Van Ryzin, G. G. & Lavena, C. F. (2013). The credibility of government performance reporting. Public Performance & Management Review, 37(1), 87-103.

Yang, K. (2009). Examining perceived honest performance reporting by public organizations: Bureaucratic politics and organizational practice. Journal of Public Administration Research and Theory, 19(1), 81-105.

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

Yoon, K., Kim, C. H. & Kim, M-S. (2011). A cross-cultural comparison of the effects of source credibility on attitudes and behavioural intentions. Mass Communication and Society, 1(3-4), 153-73.

TABLE 1. REGRESSION ANALYSIS OF MAIN AND INTERACTION EFFECTS

| | Credibilitiy index (0 to 100) | | | | Best guess true level (-3 to 3) | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Source (gov)	-2.384	0.252	-2.446	0.238	-0.057	0.605	-0.058	0.601
Agency (CIS)	-3.714	0.076	-3.558	0.088	-0.145	0.191	-0.139	0.209
Left digit (9)	0.077	0.971	0.108	0.959	-0.027	0.805	-0.027	0.809
Source * Agency	---	---	0.569	0.891	---	---	0.081	0.713
Source * Left digit	---	---	-12.845	0.002	---	---	-0.440	0.047
Agency * Left digit	---	---	-3.029	0.467	---	---	-0.120	0.589
Constant	47.338	---	47.361	---	-0.859	---	-0.858	---
R-squared	0.007		0.024		0.003		0.011	
Obs	589		589		587		587	

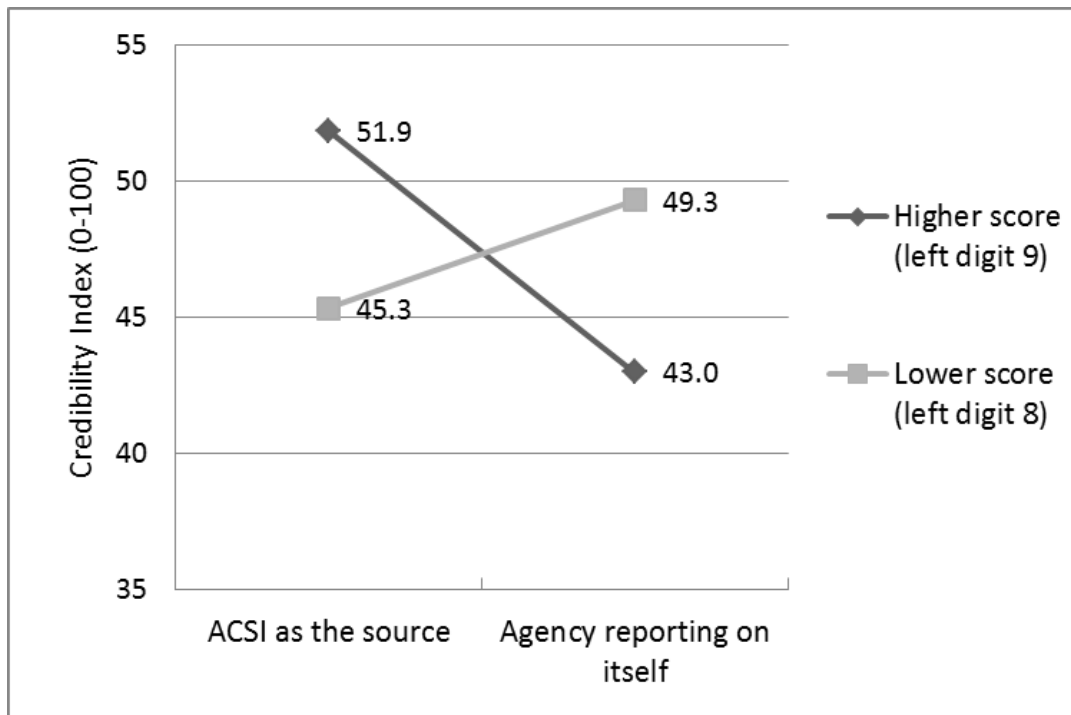Note: Table shows unstandardized coefficients from OLS regressions.

Note: Only the main effect between agencies is marginally significant (p = .076), see model 1 in Table 1.

FIGURE 1: BEST-GUESS ABOUT TRUE LEVEL (MEANS) BY TREATMENT GROUP

Note: No significant agency or source effects (see Table 1).

Note: Graph based on coefficients from Table 1, Model 2. Interaction of source by left-digit is statistically significant (p < .01).

APPENDIX A: EXPERIMENTAL DESIGN

Factor 1: Agency (randomized)

The **US Department of Veterans Affairs** is a government agency that provides services and support for people who have served in the US armed forces.

The **US Citizenship and Immigration Services** is a government agency that provides processing of visas and citizenship applications for immigrants.

Factor 2: Source (randomized)

The **US Department of Veterans Affairs / US Citizenship and Immigration Services** measures the level of customer satisfaction with its services and issues an annual report to the public. Below are the latest results (on an index where 0 is the lowest possible satisfaction and 100 is the highest possible satisfaction).

The **American Customer Satisfaction Index**, an independent non-governmental organization, measures the level of customer satisfaction with the agency's services and issues an annual report to the public. Below are the latest results (on an index where 0 is the lowest possible satisfaction and 100 is the highest possible satisfaction).

Factor 3: Index level (randomized)
Satisfaction Index: {86-93}

APPENDIX B: DEMOGRAPHIC PROFILE OF STUDY PARTICIPANTS BY EXPERIMENTAL FACTOR

	Source			Agency			Left digit	
	ACSI	Agency		VA	CIS		8_	9_
Age (years)	48.4	48.9		49.0	48.3		48.5	48.8
Sex (female)	69%	65%		67%	67%		68%	67%
Race (white)	83%	79%		84%	79%		81%	82%
Education (years)	14.7	14.7		14.6	14.8		14.7	14.7
Income (dollars)	70,269	80,684		71,230	79,155		74,847	75,683
Ideology (1=left to 10=right)	5.56	5.38		5.24	5.70		5.37	5.58
Trust of government (0=never to 100=always)	40.2	42.9		40.3	42.7		40.9	42.1

Note: No significant differences at $p < .10$, with Bonferroni adjustment.