10-15-2019

The Effects of Certain and Uncertain Incentives on Effort and Knowledge Accuracy

Thomas Jamieson

Nicholas Weller

The Effects of Certain and Uncertain Incentives on Effort and Knowledge Accuracy*

Thomas Jamieson
Assistant Professor
School of Public Administration
University of Nebraska at Omaha
&
Visiting Researcher
Department of Political Science
University of Waterloo
tjamieson@unomaha.edu

Nicholas Weller
Assistant Professor
Department of Political Science
University of California, Riverside
nweller@ucr.edu

Paper accepted for publication at the *Journal of Experimental Political Science*. Please cite as:

Jamieson, Thomas, and Nicholas Weller. 2019. "The Effects of Certain and Uncertain Incentives on Effort and Knowledge Accuracy." *Journal of Experimental Political Science*. Cambridge University Press, 1–14. https://doi.org/10.1017/XPS.2019.27.

Abstract

In many situations incentives exist to acquire knowledge and make correct political decisions. We conduct an experiment that contributes to a small but growing literature on incentives and political knowledge, testing the effect of certain and uncertain incentives on knowledge. Our experiment builds on the basic theoretical point that acquiring and using information is costly, and incentives for accurate answers will lead respondents to expend greater effort on the task and be more likely to answer knowledge questions correctly. We test the effect of certain and uncertain incentives and find that both increase effort and accuracy relative to the control condition of no incentives for accuracy. Holding constant the expected benefit of knowledge, we do not observe behavioral differences associated with the probability of earning an incentive for knowledge accuracy. These results suggest that measures of subject performance in knowledge tasks is contingent on the incentives they face. Therefore, to ensure the validity of experimental tasks and the related behavioral measures we need to ensure a correspondence between the context we are trying to learn about and our experimental design.

Keywords: incentivized survey experiment, effort, political knowledge, information search, public opinion, international affairs, predictions
Word count: 5,444

*The research design of the paper was presented at the 2017 ISA Annual Convention, at the EITM Summer Institute at the University of Houston, and in the Networked Democracy Lab at the University of Southern California. We would especially like to thank Pablo Barberá, A. Burcu Bayram, Harold Clarke, Gail Buttorff, Francisco Cantú, Dennis Chong, Douglas Dion, Nehemia Geva, Jim Granato, Patrick James, Brian Rathbun, Frank Scioli, Philip Seib, Rick Wilson, Sunny Wong, Jonathan Woon, participants in the panels, the anonymous reviewers and the Associate Editor for excellent comments and suggestions. Any errors that remain are our own responsibility. This research was supported by a USC Dornsife Gold Family Fellowship and the University of California, Riverside. The authors are aware of no conflicts of interest regarding this research. The data, code, and any additional materials required to replicate all analyses in this article are available at the Journal of Experimental Political Science Dataverse within the Harvard Dataverse Network, at: https://doi:10.7910/DVN/WVFZGE.

Many political decisions happen in a context in which individuals face incentives to acquire information and use it to make an inference about a future event. For example, voters have an incentive to select the candidate that best represents their interests and desired policies, which might be affected by knowledge about foreign trade, immigration, or international conflict. Voters would therefore have an incentive to acquire information about these topics and use it to help determine which candidate to support. In this simple example, which occurs regularly in politics, there are incentives for accurate knowledge of the world, and yet few experiments have investigated how incentives affect political knowledge. If incentives matter in these situations and we fail to study their effects, then we may misunderstand behavior in many political contexts.

Prior research about political knowledge and behavior demonstrates that incentives can systematically affect people's knowledge about domestic politics (Krupnikov et al., 2006; Prior and Lupia, 2008). We extend research about incentives and knowledge by examining how uncertainty in incentives affects knowledge. To do so, we examine knowledge questions for which the outcomes have not yet occurred, and therefore answering these questions correctly requires a combination of information about the current state of the world and the ability to make an inference about how things will change five weeks into the future. We show that incentives increase respondents' effort and accuracy in answering knowledge questions, and subjects behave similarly whether the incentives are guaranteed or uncertain.

This paper contributes to the methodological literature focused on experimental design (Morton and Williams, 2010). Good experimental design captures the essential elements of the theory being tested, and if the underlying theory involves individuals having an incentive for accurate knowledge or information acquisition, then our experiments need to include similar incentives. In their absence, an experiment will not be a good match to the behavior being studied.

1 Incentives, Effort, and Knowledge Accuracy

In this paper we examine the ability of voters to identify the correct answer to a question where the outcome will not be known until a future, but not very distant, date. The ability to predict the outcome of an event in which initial conditions are knowable but there is some uncertainty about the future is a common, politically-relevant task (Lupia and McCubbins, 1998). Research demonstrates

that incentives for accuracy improve the ability of people to correctly answer political knowledge questions (Feldman, Huddy and Marcus, 2015; Prior and Lupia, 2008), and financial incentives encourage subjects to update their beliefs about political facts (Hill, 2017).

We contribute to the literature on incentives and knowledge by studying both certain and uncertain incentives. Research has generally used incentives where the payoffs for correct knowledge occur with certainty; however, Hill (2017) does examine how probabilistic incentives affect learning. In many political contexts incentives are uncertain and therefore understanding their effects is important to learning about behavior. For example, acquiring accurate knowledge may not change one's vote with certainty, and even if it does affect a vote, the election outcome and its benefits are uncertain. In either case the benefit of correct knowledge is uncertain and this uncertainty may affect behavior.[1] As elaborated below, we expect both certain and uncertain incentives to improve knowledge accuracy because they both affect the expected benefits of knowledge, which should lead to more correct answers.

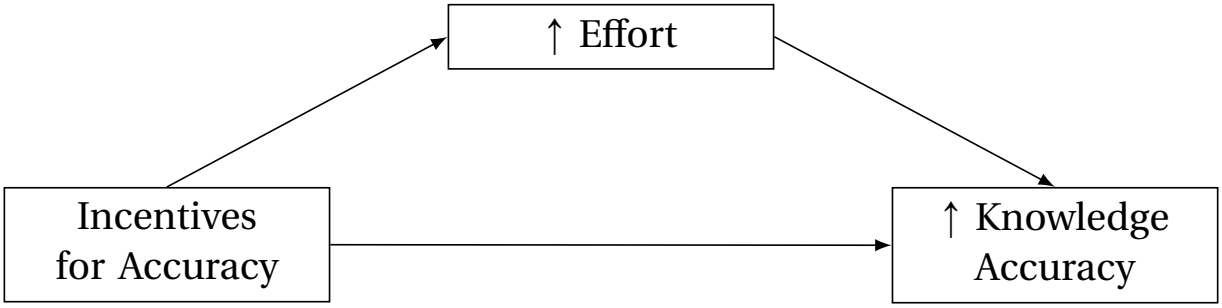
Prior research has also focused on situations in which the correct answer to a question can be identified relatively easily online or in a book. However, the questions we use in our experiment (described in the next section) ask respondents to identify the right answer to a question when the outcome will not be known for about five weeks. Furthermore, we focus on knowledge in the realm of foreign affairs. Prior research suggests that while it is generally difficult to identify future outcomes in international affairs (Tetlock, 1998, 1999, 2006), some people appear able to make accurate predictions (Mellers et al., 2015; Tetlock, 1998, 1992; Tetlock and Gardner, 2016). We do not look at individual-level factors that correlate with prediction ability, and instead we focus on whether incentives for accuracy improve people's ability to make accurate judgments about future outcomes.

In many ways this is a hard test of responses to financial incentives given that people's knowledge of and interest in international affairs typically lags behind domestic politics (Converse, 1964; Delli Carpini and Keeter, 1996; Holsti, 1992; Kinder and Sears, 1985; Lupia, 2015). However, foreign policies can have significant effects on people's lives and well-being. For example, military

[1] In other contexts, some research suggests that participants are more responsive to guaranteed incentives than uncertain payoffs, even if the expected utility is held constant (Pforr et al., 2015; Warriner et al., 1996; Zheng, Gong and Pavlou, 2017). However, these other studies have focused on domains very different than ones in which the incentives increase the potential benefit of cognitive effort and knowledge.

conflict, international agreements, international trade policies, immigration policies, membership of international organizations, and economic integration have widespread implications for the economic prosperity and security of domestic populations. The importance of international affairs was highlighted in the 2016 U.S. Presidential election, where immigration and globalization played a significant role in the campaign.

Figure 1: A Theory of Incentives, Effort, and Knowledge Accuracy.



Our basic model of how incentives affect knowledge accuracy is displayed in Figure 1. In the rest of this section we elaborate on both the direct path by which incentives improve accuracy and the indirect path in which increased effort leads to an increase in knowledge accuracy.

1.1 Effect of Incentives on Effort

Answering questions correctly requires that respondents pay attention to the task and expend cognitive effort to acquire and use knowledge in making predictions about future events (Lupia and McCubbins, 1998). Individuals will be more likely to incur these costs if there are benefits for doing so, and we expect individuals to expend greater effort to make accurate judgments when the expected benefit of effort increases.

Hypothesis 1 *Incentives and Effort Hypothesis*

If there are incentives for accuracy, then individuals are more likely to expend effort than in the absence of incentives.

In our experiment we provide a financial incentive for correct answers, but outside of the experimental context incentives could be any factors that increase the value of accurate knowledge.

1.2 Effect of Incentives on Knowledge Accuracy

Prior studies show that people respond to financial incentives with improved political knowledge (Prior and Lupia, 2008); people update their beliefs in response to incentives, even if they are not perfect Bayesians (Hill, 2017); and incentives for accuracy reduce partisan bias (Prior, Sood and Khanna, 2015). We expect that incentives lead to improved knowledge accuracy through both the direct effect of incentives and an indirect effect of greater effort. This leads to the second set of hypotheses:

Hypothesis 2 *Incentives and Knowledge Hypothesis*

If there are incentives for accuracy, then individuals are more likely to correctly answer questions than if such incentives are absent.

Hypothesis 3 *Incentives and Knowledge Mediation Hypothesis*

If there are incentives for accuracy, then individuals are more likely to expend effort than when there is not an incentive for accuracy and effort will lead to improved accuracy.

Incentives should cause participants to engage in greater effort, and therefore improve knowledge accuracy compared to participants in the control condition.

2 Experimental Design

We recruited 1016 subjects using Amazon MTurk for the experiment, which we designed to isolate the effects of incentives on effort and accuracy of answers to political knowledge questions. Previous research suggests that the MTurk platform produces acceptable samples for social science research (Berinsky, Huber and Lenz, 2012; Casler, Bickel and Hackett, 2013; Hauser and Schwarz, 2016; Huff and Tingley, 2015; Levay, Freese and Druckman, 2016; Mullinix et al., 2015).[2] We expect our theory about incentives, effort, and knowledge accuracy to apply to all people, and therefore there is no reason to expect MTurk respondents to be theoretically inappropriate for our purposes.

[2]After an initial pilot study, we implemented an improved experimental design to test the relationship between incentives, effort, and knowledge accuracy. Further details about the design and the results of the pilot are reported in Appendix H. We thank the anonymous reviewers for their suggestions for improvements in the design.

The additional advantage of MTurk is that we know subjects are at an internet-connected device and can search for information, and the platform also provides a way to track how long respondents take to complete the task, which we can use as a proxy measure for effort. Appendix C presents descriptive statistics of the sample.

After being recruited via MTurk, but prior to treatment assignment, participants received background information about the experiment and they answered a series of political information questions. We used ten political information questions that covered a mix of U.S. domestic issues and foreign affairs; the questions are displayed in section 1.2.2 in Appendix B. After answering these questions, respondents were randomly assigned to the control group or one of three treatment groups.

In Table 1 we report the average number of correct answers to the pre-treatment political information questions for respondents assigned to each treatment condition. As expected, the averages do not vary across treatment conditions. We measured pre-existing political information so that we can use it to contextualize the magnitude of estimated treatment effects by comparing them to the relationship between prior knowledge and the questions we use as our dependent variable.

Table 1: Descriptive Statistics of Pre-treatment Political Information, by Treatment.

	Count	Mean	Std. Dev.	Min.	Max.	T-test P-Value
Control Condition	256	3.707	1.416	1	8	
Bonus Treatment	256	3.543	1.316	1	7	0.175
Random Bonus Treatment	250	3.540	1.383	1	8	0.180
Lottery Treatment	254	3.531	1.315	1	7	0.148
Total	1016	3.581	1.358	1	8	

The control group received no incentives for accurate answers. We used three different incentive schemes in the experiment and held constant the expected value of a right answer to an outcome question at \$0.50. We varied the benefit of a correct answer and the uncertainty associated with receiving the benefit if a question was answered correctly.

In the bonus treatment participants received a \$0.50 guaranteed payment for every correct answer. In the random bonus treatment we randomly selected one of the five questions and paid respondents \$2.50 if they got that particular question correct. In the lottery treatment each participants earned a ticket for a \$50 lottery for each correct answer,and one \$50 prize was awarded

for every 100 correct answers.

We asked comprehension questions about the incentives immediately after exposure to the treatment but prior to our outcome questions. This helped ensure that participants understood the bonuses and their likelihood of receiving a bonus given a certain score (Kane and Barabas, N.d.).

To identify the effect of incentives on effort and knowledge accuracy we chose to examine whether respondents got the right answer to five different questions that asked about an outcome that would not be officially known for about five weeks after the respondents were asked to predict the answer to the question.[3] The values for all five questions changed over the five weeks, but because of lumpiness in the response categories the correct answer changed for only three questions over the time period of the study.

All questions were multiple choice and the survey required participants to provide a response for each question to continue the survey. The number of possible responses ranged from six to eleven responses, which varied according to plausible answers to each question. Table 2 presents the full text of each question, including the number of possible responses. The entire question wording and possible responses are available to view in Appendix B.

In all of the conditions participants were allowed and encouraged to search for information to answer the questions. Making this consistent across conditions allows us to minimize the possibility that subjects in the treatment conditions inferred that we wanted them to search for information whereas those in the control condition might not make that inference, and therefore our treatment effects would be confounded by subjects' perceptions of what is expected of them.

These outcome questions represent a combination of correct information and the ability to use it to answer a question about the near-term future. We refer to the ability to answer such question as knowledge following the distinction between information as data and knowledge as the ability to make accurate predictions (Lupia and McCubbins, 1998).

Because the correct answer would not be known for five weeks, respondents cannot simply look up the answers, but the outcomes also do not occur as far into the future as many of the predictions used in the Good Judgment Project or other forecasting examples.

[3]We chose five weeks for both design-based and logistical reasons. We wanted enough time that the current value might change, but at the same time we needed the timeline to be short enough that respondents would get paid soon enough to take the incentive seriously.

Table 2: Knowledge Accuracy Questions and Correct Answers.

No.	Topic	Question	Choices	Correct Answer
1	U.S. Jihadist Attacks	According to New America, a non-partisan think tank, what will be the figure CLOSEST to the number of Jihadist terrorist attacks in the U.S. committed by people who were U.S. CITIZENS OR PERMANENT RESIDENTS at the time of charge or death from January 1, 2018 to November 30, 2018?	8	8
2	Syria Civilian Deaths	According to IAmSyria.org, a non-profit campaign, what will be the figure CLOSEST to the number of CIVILIANS KILLED IN THE SYRIA CONFLICT from January 1, 2018 to November 30, 2018?	6	6,500
3	Mueller Indictments	Special Counsel for the U.S. Department of Justice, Robert Mueller, is leading an investigation into Russian interference in the 2016 Presidential election. What is the figure CLOSEST to the NUMBER OF PEOPLE INDICTED OR GIVEN PLEA DEALS in the investigation as of November 30, 2018?	8	30
4	U.S. Refugee Resettlements	According to the U.S. Refugee Processing Center, what will be the figure CLOSEST to the NUMBER OF REFUGEES RESETTLED IN THE U.S. from January 1, 2018 to November 30, 2018?	11	20,000
5	U.S. Deaths in Afghanistan	According to iCasualties.org, an independent website to track casualties, what will be the figure CLOSEST to the number of U.S. MILITARY DEATHS in and around AFGHANISTAN from January 1, 2018 to November 30, 2018?	6	10

For Question 1, a participant could search for information about the number of Jihadist terrorist attacks in the U.S. committed by people who were U.S. citizens or permanent residents at the time of the experiment, and this would help them answer the question correctly. Some of the possible answers were either impossible or highly unlikely given the prior number of terrorist attacks, but a correct answer still requires making a judgment about the number of events in the five weeks before correct answers were determined. An accurate response may reflect a combination of information search and the ability to make an inference about how the values will change over the course of five weeks.[4]

The short time between asking the outcome questions and paying subjects means that the correct answers to our questions will necessarily be relatively close to the state of the world when taking the survey. However, the use of these kinds of questions are not wholly different than some used in other studies of predictions. For instance, (Tetlock and Gardner, 2016, 125-126) discuss the following example from their work:

> As the Syrian civil war raged, displacing civilians in vast numbers, the IARPA tournament asked forecasters whether "the number of registered Syrian refugees reported by the United Nations Refugee Agency as of 1 April 2014" would be under 2.6 million. That question was asked in the first week of January 2014, so forecasters had to look three months in the future.

Like our questions, this one requires some information about current conditions and the ability to predict into the short-term future, which requires consideration of the trends underlying the current state of the world and projecting them forward. The questions we use are neither as easy as straightforward information questions nor as difficult to answer correctly as predictions with an 18-month time frame.

After the knowledge questions, participants were asked a series of questions relating to effort and information search.[5] The full design, including the treatments, questions, and coding of variables is presented in Appendix B.

We used two attention checks during the experiment to ensure our subjects were attentive.

[4] We thank an anonymous reviewer for this well-made point.

[5] The experiment also featured batteries of political attitude questions, and the results of these are presented in other papers.

Almost ninety percent of the sample answered both questions correctly.[6]

The experiment was launched on October 26, 2018 and participants were paid bonuses on December 1, 2018.

3 Results

In this section, we discuss the results of our experiments. Overall, we find that incentives increase respondent effort and improve the accuracy of answers to the questions. The effect of incentives does not appear to vary with their uncertainty.

3.1 Incentives and Effort

Incentives increase the amount of effort expended. Pooled across all treatments, incentives for accuracy increased effort, measured as time spent on the survey, from a baseline of 10.83 minutes in the control group to 11.77 minutes across all treatments, an 8.71 percent increase in time spent. In both the control and treatment groups, subjects were told that they could use the internet to help them answer the questions correctly so this increase in effort is solely due to the incentives for accuracy.

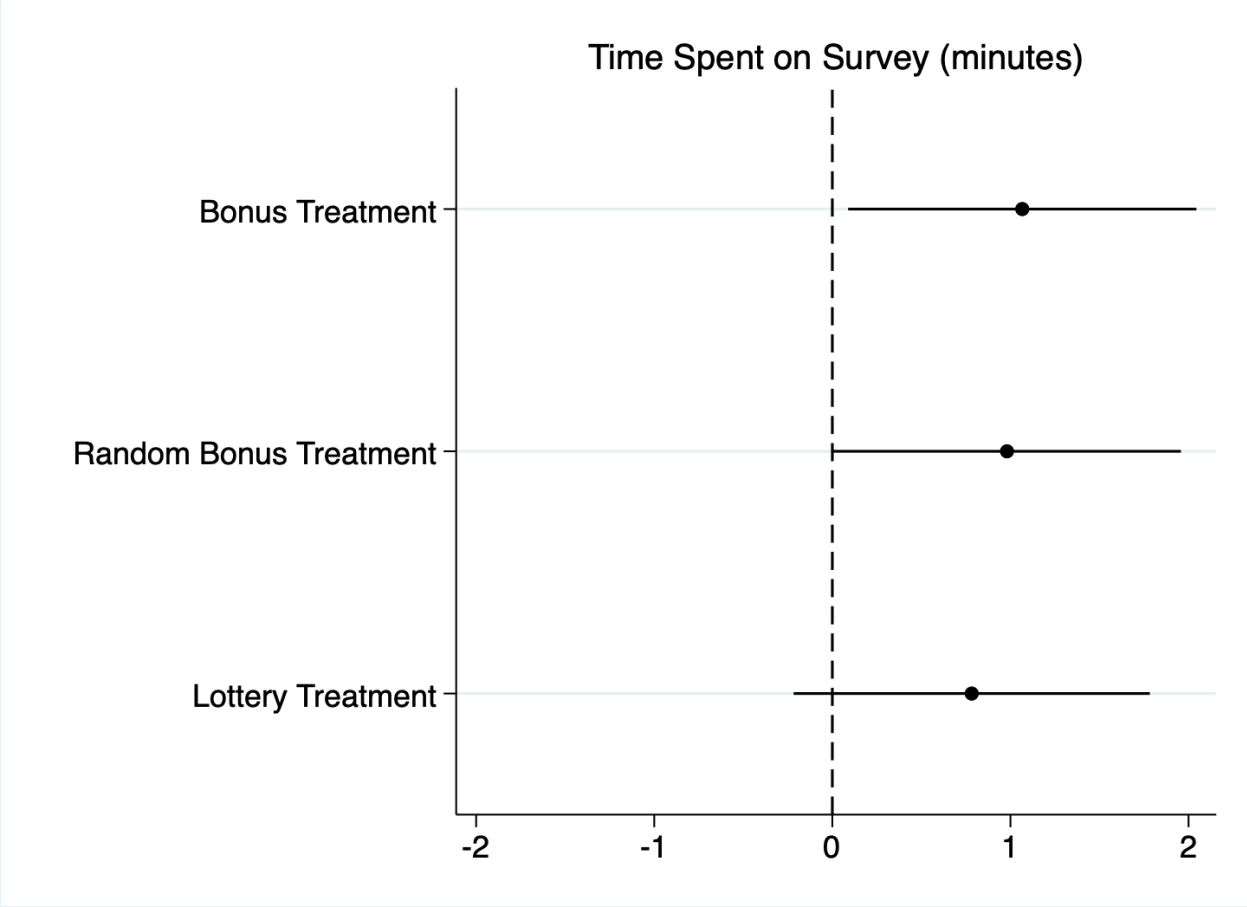
Figure 2 illustrates the effect of each individual treatment on effort. The point estimates for the three treatments are indistinguishable from one another; although the lottery treatment just misses standard statistical significance levels (p=0.125). These results are consistent with our expectations by demonstrating that even small incentives increase effort. Furthermore, as expected the increase in effort does not vary with uncertainty because the expected value of effort is equivalent across treatments.

3.2 Incentives and Knowledge Accuracy

The experimental results also demonstrate that incentives for accuracy increase the number of knowledge questions answered correctly. Table 3 presents descriptive statistics for knowledge accuracy across the experimental conditions. Pooled across all treatments, incentives for accuracy improved the number of correct answers by 0.28. On average subjects in the control condition

[6] As suggested by Berinsky, Margolis and Sances (2014), we did not drop participants if they failed the screening questions, but we report the results in Appendix D.

Figure 2: Determinants of Effort.
Dependent Variable: Time Spent on Survey (minutes).



Point estimates indicate time spent with 95% confidence intervals using `coefplot` in `Stata` (Jann, 2014).

answered less than one of the five knowledge questions correctly. The modal number of correct answers was zero in the control group, and it was one in the treatment groups. So, while some information about the right answers was available online, it is clear that subjects still had a hard time answering the questions correctly.

Table 3: Descriptive Statistics of Predictive Accuracy, by Treatment.

	Count	Mean	Std. Dev.	Min.	Max.	T-test P-Value
Control Condition	256	0.797	0.889	0	5	
Bonus Treatment	256	1.172	1.018	0	4	0.000
Random Bonus Treatment	250	0.964	0.995	0	5	0.047
Lottery Treatment	254	1.091	0.996	0	5	0.001
Total	1016	1.006	0.985	0	5	

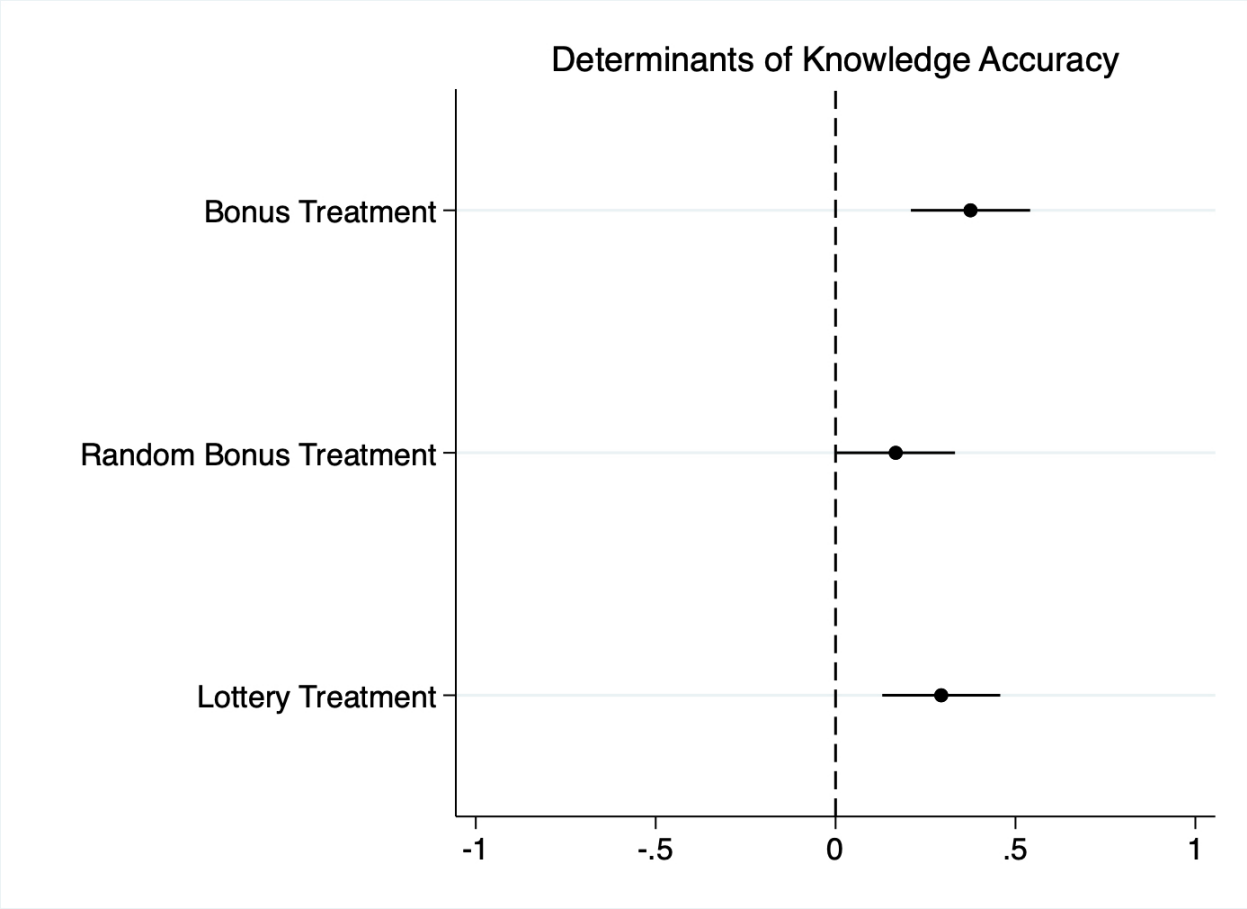
All three incentive conditions led to increased accuracy compared to the control condition. The bonus treatment increased accuracy by 0.375 correct answers, the random bonus treatment increased accuracy by 0.167 correct answers, and the lottery treatment increased accuracy by 0.29 correct answers.

Furthermore, the treatment effects are indistinguishable from each other, suggesting they have the same average effect on behavior. Significantly, Figure 3 illustrates the absence of differences between treatments on knowledge accuracy — all three incentives led to improved performance.

To put these effect sizes in perspective, we compare the magnitude of the average treatment effect across all three conditions to the estimated magnitude of the relationship between pre-treatment political information and accurate knowledge in the control group. Our regression estimates indicate that each additional pre-treatment political information question answered correctly increases by 0.07 the number of knowledge questions answered correctly. Recall that the average treatment effect across incentive conditions is an increase of 0.28 in the number of correct answers, which is equivalent to moving from a respondent with an average level of baseline information to a respondent with the highest observed level of baseline political information in our control group.

The results indicate that we may understate both respondents' willingness to expend effort on a task, and their ability to correctly answer knowledge questions if we fail to implement the incentive conditions that best match the theoretical or real-world context being studied.

Figure 3: Determinants of Knowledge Accuracy.
Dependent Variable: Number of Correct Answers.



Point estimates indicate the number of correct answers with 95% confidence intervals using `coefplot` in `Stata` (Jann, 2014).

3.3 Incentives, Effort, and Knowledge Accuracy

To determine whether effort mediates the effects of the treatments on the participants' knowledge accuracy, we use causal mediation analysis (Imai et al., 2011; Imai, Keele and Yamamoto, 2010).[7]

Table 4: Direct Treatment Effects and Mediation Effects of Incentives on Knowledge Accuracy. *Mediating Variable: Time Spent on Survey (minutes).*

	ACME	Direct Effect	Total Effect	Prop. Total Effect Mediated
Bonus Treatment	0.053	0.328	0.382	0.139
	(0.005, 0.107)	(0.168, 0.484)	(0.225, 0.549)	(0.097, 0.237)
Random Bonus Treatment	0.034	0.135	0.170	0.198
	(-0.001, 0.076)	(-0.026, 0.292)	(0.013, 0.336)	(0.092, 1.246)
Lottery Treatment	0.041	0.271	0.312	0.131
	(-0.009, 0 .094)	(0.115, 0.423)	(0.157, 0.475)	(0.087, 0.262)

The results were calculated using 1000 simulations with 95% confidence intervals in brackets using `mediation` in `Stata` (Hicks and Tingley, 2011).

Table 4 presents the results of causal mediation analysis. Across all treatments, effort appears to be an important mediator of correct answers. Effort accounts for almost 14 percent of the average treatment effect of the bonus treatment, nearly 20 percent of the random bonus treatment, and about 13 percent of the lottery treatment on knowledge accuracy. All three of these estimates are of similar magnitude, suggesting that the effort induced by the incentives has a similar mediating effect on knowledge accuracy.

3.3.1 Text Analysis

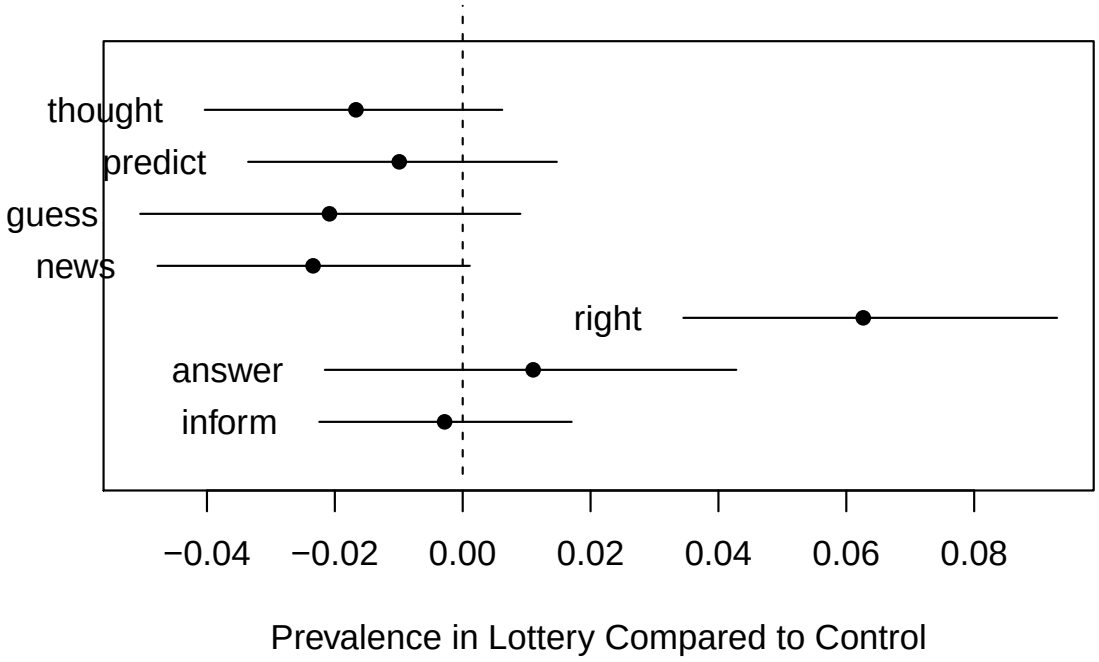
To further explore the mechanisms influencing knowledge accuracy, we use the structural topic model `stm` in `R` developed by Roberts et al. (2014). Similar to Mildenberger and Tingley (2017), we examine respondents' responses to an open-ended question about their thoughts as they made answered the knowledge questions.[8] We identified that seven topics was an appropriate number for subjects' responses in the experiment.

We then estimated the difference in the prevalence of each topics between the treatment conditions and the control condition. In comparing each treatment to the control, we consistently found that only one topic was consistently more common in the treatment than the control, and this topic was associated with words related to answering questions correctly.

[7] Appendix F presents the full results of this analysis.

[8] A full discussion of these methods and results are included in Appendix G.

Figure 4: Effect of Lottery on Free Responses by Respondents



Point estimates and 95% confidence intervals. The plot was created using the `stm` package in `R` (Roberts, Stewart and Tingley, 2018).

This result is presented visually in Figure 4 in which we plot the seven topics and the estimated difference for each topic between the lottery condition and the control condition; the topic labels are based on the most commonly-appearing words for each of the seven topics. The results are substantively similar for both of the other two treatments.

The text analysis provides further evidence that incentives affect behavior and that the different incentives are broadly similar in their effects.

4 Discussion

In this paper, we demonstrate that incentives for accuracy increase both effort and the number of correct answers to knowledge questions about international affairs. Furthermore, our experiment shows that behavior does not vary with the uncertainty of the accuracy incentive. The results provide evidence for the importance of understanding the context in which political decisions are made and have both substantive and methodological importance for political science.

Substantively, individuals may be more capable of understanding and reaching accurate answers about politics than often found in previous research that uses non-incentivized behavior. Our results show that measures and estimates of citizen knowledge may be affected by context. People increase their effort and knowledge about international affairs when given incentives for being correct. Given the stakes of political decisions (Edlin, Gelman and Kaplan, 2007), there can be quite large incentives to correctly understand political outcomes. Even though incentives exist in real political decisions, incentives for accuracy often do not exist in many experiments or survey settings. The results suggest that when looking at political behavior, we should consider the extent to which there were incentives for accuracy and how their absence or presence affects our interpretation of the observed behavior.

The possible mismatch between experimental/survey design and our theories is important for political behavior scholarship. Beyond the risks of "cheating" in these batteries of knowledge (Barabas et al., 2014; Clifford and Jerit, 2016), the absence of explicit incentives means our surveys/experiments may not offer appropriate insight into situations in which there is utility associated with making the correct decision. For example, if an experiment focuses on how knowledge affects voting decisions, then we need to ensure that the experimental context captures the incen-

tives for accurate knowledge at the ballot box.

Our experiment and results also suggest directions for future studies. First, future research could introduce explicit costs to searching for information so that we have a better sense of how the cost of effort affects behavior or measure effort directly through unobtrusive observations of individuals' online search behavior and attention to international affairs. Second, the relatively small incentives in this study could mean we underestimate incentives' effect. Future studies could test the effects of larger incentives on individual behavior. Third, incentives to make accurate judgments may be framed negatively as costs for making mistakes, which could produce systematic differences in individual behavior (Tversky and Kahneman, 1981). The general point is that we still have much to learn about how costs and incentives affect effort, knowledge acquisition and knowledge accuracy.

References

Barabas, Jason, Jennifer Jerit, William Pollock and Carlisle Rainey. 2014. "The Question(s) of Political Knowledge." *American Political Science Review* 108(4):840–855.

Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–368.

Berinsky, Adam J., Michele F. Margolis and Michael W. Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 58(3):739?753.

Casler, Krista, Lydia Bickel and Elizabeth Hackett. 2013. "Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing." *Computers in Human Behavior* 29(6):2156–2160.

Clifford, Scott and Jennifer Jerit. 2016. "Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions." *Public Opinion Quarterly* 80(4):858–887.

Converse, Philip E. 1964. "The nature of belief systems in mass publics." *Critical Review* 18(1-3):1–74.

Delli Carpini, Michael X. and Scott Keeter. 1996. *What Americans Know about Politics and why it Matters*. Yale University Press.

Edlin, Aaron, Andrew Gelman and Noah Kaplan. 2007. "Voting as a Rational Choice: Why and How People Vote To Improve the Well-Being of Others." *Rationality and Society* 19(3):293–314.

Feldman, Stanley, Leonie Huddy and George E. Marcus. 2015. *Going to War in Iraq: When Citizens and the Press Matter*. University of Chicago Press.

Hauser, David J. and Norbert Schwarz. 2016. "Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants." *Behavior Research Methods* 48(1):400–407.

Hicks, Raymond and Dustin Tingley. 2011. "Causal mediation analysis." *Stata Journal* 11(4):605–619.

Hill, Seth J. 2017. "Learning together slowly: Bayesian learning about political facts." *Journal of Politics* 79(4).

Holsti, Ole R. 1992. "Public Opinion and Foreign Policy: Challenges to the Almond-Lippmann Consensus." *International Studies Quarterly* 36(4):439–466.

Huff, Connor and Dustin Tingley. 2015. ""Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents." *Research & Politics* 2(3).

Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4):765–789.

Imai, Kosuke, Luke Keele and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1):51–71.

Jamieson, Thomas and Nicholas Weller. 2019. "Replication Data for: The Effects of Certain and Uncertain Incentives on Effort and Knowledge Accuracy." *Harvard Dataverse,* V1.
URL: *doi:10.7910/DVN/WVFZGE*

Jann, Ben. 2014. "Plotting Regression Coefficients and Other Estimates." *The Stata Journal* 14(4):708–737.

Kane, John V. and Jason Barabas. N.d. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science*. Forthcoming.

Kinder, Donald R. and David O. Sears. 1985. "Public opinion and political action." *Handbook of Social Psychology* 2:659–741.

Krupnikov, Yanna, Adam Seth Levine, Arthur Lupia and Markus Prior. 2006. "Public Ignorance and Estate Tax Repeal: The Effect of Partisan Differences and Survey Incentives." *National Tax Journal* 59(3):425–437.
URL: *https://www.jstor.org/stable/41790333*

Levay, Kevin E., Jeremy Freese and James N. Druckman. 2016. "The Demographic and Political Composition of Mechanical Turk Samples." *SAGE Open* 6(1).

Lupia, Arthur. 2015. *Uninformed: Why People Seem to Know So Little about Politics and What We Can Do about It*. New York: Oxford University Press.

Lupia, Arthur and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge, U.K. ; New York: Cambridge University Press.

Mellers, Barbara, Eric Stone, Pavel Atanasov, Nick Rohrbaugh, S. Emlen Metz, Lyle Ungar, Michael M. Bishop, Michael Horowitz, Ed Merkle and Philip Tetlock. 2015. "The psychology of intelligence analysis: drivers of prediction accuracy in world politics." *Journal of Experimental Psychology. Applied* 21(1):1–14.

Mildenberger, Matto and Dustin Tingley. 2017. "Beliefs about Climate Beliefs: The Importance of Second-Order Opinions for Climate Politics." *British Journal of Political Science* pp. 1–29.

Morton, Rebecca B. and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality*. New York: Cambridge University Press.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2(2):109–138.

Pforr, Klaus, Michael Blohm, Annelies G. Blom, Barbara Erdel, Barbara Felderer, Mathis Fräßdorf, Kristin Hajek, Susanne Helmschrott, Corinna Kleinert, Achim Koch, Ulrich Krieger, Martin Kroh, Silke Martin, Denise Saßenroth, Claudia Schmiedeberg, Eva-Maria Trüdinger and Beatrice Rammstedt. 2015. "Are Incentive Effects on Response Rates and Nonresponse Bias in Large-scale, Face-to-face Surveys Generalizable to Germany? Evidence from Ten Experiments." *Public Opinion Quarterly* 79(3):740–768.

Prior, Markus and Arthur Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American Journal of Political Science* 52(1):169–183.

Prior, Markus, Gaurav Sood and Kabir Khanna. 2015. "You Cannot be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10(4):489?518.

Roberts, Margaret E., Brandon M. Stewart and Dustin Tingley. 2018. *stm: R Package for Structural Topic Models*. R package version 1.3.3.
URL: *http://www.structuraltopicmodel.com*

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–1082.

Tetlock, Philip. 1998. "Close-Call Counterfactuals and Belief-System Defenses: I Was Not Almost Wrong but I Was Almost Right." *Journal of Personality and Social Psychology* 75:639–652.

Tetlock, Philip E. 1992. "Good Judgment in International Politics: Three Psychological Perspectives." *Political Psychology* 13(3):517–539.

Tetlock, Philip E. 1999. "Theory-Driven Reasoning About Plausible Pasts and Probable Futures in World Politics: Are We Prisoners of Our Preconceptions?" *American Journal of Political Science* 43(2):335–366.

Tetlock, Philip E. 2006. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

Tetlock, Philip E and Dan Gardner. 2016. *Superforecasting: The art and science of prediction.*

Random House.

Tversky, A. and D. Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science (New York, N.Y.)* 211(4481):453–458.

Warriner, Keith, John Goyder, Heidi Gjertsen, Paula Hohner and Kathleen Mcspurren. 1996. "Charities, No; Lotteries, No; Cash, Yes: Main Effects and Interactions in a Canadian Incentives Experiment." *Public Opinion Quarterly* 60(4):542–562.

Zheng, Alvin, Jing Gong and Paul Pavlou. 2017. "On Using the Lottery in Crowdfunding Platforms: 'Crowding in' the Masses But 'Crowding out' Success.".
URL: *https://papers.ssrn.com/abstract=2916807*