# What Citizens Know Depends on How You Ask Them:

# Experiments on Time, Money and Political Knowledge

Markus Prior, Princeton University

Arthur Lupia, University of Michigan

**Version 2.0. October 2005.**

We thank the Center for Political Studies at the University of Michigan and the University Committee on Research in the Humanities and Social Sciences at Princeton University for funding this research. We thank Rick Li, Mike Dennis, Bill McCready and Vicki Huggins at Knowledge Networks for assistance in programming and implementing the study. We thank Larry Bartels, James Druckman, Elisabeth Gerber, Jon Krosnick, Adam Levine, Jesse Menning, Norbert Schwarz, and seminar participants at the Midwest Political Science Association, the American Political Science Association and Princeton University for helpful advice.

**Abstract**

Surveys provide widely cited measures of political knowledge. Do unusual aspects of survey interviews affect these measures? An experiment on a nationally representative sample of over 1200 Americans provides an answer. Respondents are randomly assigned to one of four groups. A control group answers questions in a typical survey context. Respondents in three treatment groups are given a longer window of time in which to answer questions, a small monetary incentive for answering questions correctly, or both. These variations increase performance significantly for almost every knowledge question we asked. Overall, average knowledge scores in the treatment groups are 11-24 percent higher than in the control group. The treatments also cause significant reductions in the magnitude of respondents' errors on open-ended questions. The findings imply that new elicitation strategies can improve our understanding of what citizens know about politics and other socially relevant phenomena.

A basic premise of democratic governance is that citizens know certain things about politics. For this reason, scholars and other political observers devote considerable energy to the study of political knowledge, relying mostly on large surveys for evidence. These data have yielded two conclusions. First, many citizens do badly at answering political knowledge questions (e.g., Converse 1964, Kinder and Sears 1985, Bennett 1995). Second, most analysts see such performances as a problem for democratic governance. The public's inability to answer survey-based political knowledge questions is often interpreted as evidence that citizens' abilities at key democratic moments, such as elections, are far less than they could be.

While it may seem reasonable to draw conclusions about a person's ability from such responses, in other contexts this kind of inference can backfire. To see how, consider a simple example: "Professor, what percentage of the vote did John Kerry actually receive in Kansas?" Such questions from an eager undergraduate can strike fear into the heart of many lecturers. Few political scientists can answer such questions when they are asked without warning. While many scholars know where and how to find the answers, and would do so quickly if given an opportunity, the normal pace of a classroom lecture usually precludes halting the interaction to consult trusted references. In such cases, mumbling something about "a book on my shelf" or "a website that has the answer" is the best one can do from the lectern. While most people would consider it unfair for students to base broad judgments of a professor's competence on his or her immediate responses in such circumstances, common evaluations of citizens' abilities rest on just this kind of inference.

The most widely-used measures of political knowledge come from respondents' answers to survey questions. Some survey firms give respondents advance warning (e.g., a letter in the mail) that they will be contacted. Many other firms give no such notice. Of those respondents

who receive advance notice, few—if any—are given detailed information about the content of the survey. And few, if any, are told that the survey will include questions about factual aspects of politics. Survey-based political knowledge questions catch respondents by surprise.

In many cases where citizens make important political decisions, such as whether to turn out for a particular election or choose which candidate to vote for, circumstances are different. The need to make such decisions does not appear "out of the blue." Election dates, for example, are usually knowable for months or even years in advance. In such contexts, citizens are much more likely to have advance warning of a need to access particular political information than do respondents in political surveys.

Moreover, the pace of a survey interview prevents respondents from behaving as at least some of them would before key political moments. In such cases, citizens can read newspapers, visit websites, or consult with friends about important political topics. While not all citizens avail themselves of such opportunities, at least the opportunities are there to be had. The typical survey interview does not facilitate such opportunities. Their pace is established in part by conversational norms (Schwarz 1996, Chapter 5) and in part by the incentives of the interviewer (Kenickell 2000, 2003) and the respondent (Krosnick and Alwin 1987, Blair and Burton 1987). Such dynamics can lead respondents to satisfice—to offer answers without thinking hard about them—and they can induce interviewers to keep the interview moving as respondent participation is difficult to maintain in telephone interviews that go on for too long. This context does not give respondents opportunities to check the same kinds of references that at least some would use if they knew that a key democratic moment were approaching.

Differences between common attributes of the survey context and attributes of the contexts in which people make important political decisions matter because opportunity and motivation

should affect respondents' abilities to answer political knowledge questions (Luskin 1987, Delli Carpini and Keeter 1996). The difference between what we might say readily during an unexpected and fast-paced survey interview and the knowledge we need to accomplish a large range of less frequent tasks is why we keep books on our shelves, the phone numbers of trusted friends in an easily accessible place, and computers on our desks. Knowing where to find these references expands the knowledge base upon which we base important decisions. We use these resources to organize information for quick retrieval later on. Most political surveys offer neither the opportunity nor the motivation to consult such references.

In this paper, we contend that the practice of drawing broad conclusions about citizens' abilities on the basis of conventional survey-based political knowledge measures is of limited value. If the rationale for devoting so much attention and energy to questions of political knowledge is to assess the extent to which citizens can make important decisions at key democratic moments, then our elicitation procedures should better match the conditions under which they make such decisions. In what follows, we explore how inferences about political knowledge change when new standards are used. Our analysis focuses on the following claim:

> ***Null hypothesis:*** *Allowing extra time or providing a small monetary incentive for correctly answering knowledge questions during a survey interview will affect neither the likelihood of offering a correct answer nor the nature of incorrect responses.*

We use a 2x2 experimental design to test this hypothesis. Within a set of survey interviews, we give one randomly selected half of our sample only one minute to answer each question, whereas the other half has a broader time window—up to 24 hours—to respond. Independently, we offer one randomly selected half of our sample a small monetary reward ($1) for each correct answer. Each respondent is thus equally likely to be in one of four conditions: one minute no pay; one minute with pay, 24 hours no pay, 24 hours with pay. These variations are important

because the standard way of measuring political knowledge is to place all respondents in the "one minute no pay" context.[1] In such surveys, the extent to which answers to knowledge questions are a result of inhibitory attributes of common survey contexts cannot be evaluated.

As a result, the null hypothesis is far from trivial—it is a foundation of the existing literature. Survey-based claims about political knowledge presume that observed responses are robust to alterations in the survey context. Conclusions that poor performance on surveys reflects how citizens will act at key democratic moments presume that the responses are the same in the surveys as they would be at the key moments. Our work reveals that the presumptions do not hold when the survey context is altered. It also provides a constructive alternative for better conceptualizing and measuring political knowledge.

We conclude that current measures of political knowledge do not capture adequately what people know when it matters most from a democratic perspective. We reject the null hypothesis not only for the sum of correct answers to all the knowledge questions we asked but also for almost every single question when analyzed individually. On average, extra time and a small monetary incentive lead to a 24 percent increase in the number of questions answered correctly. Simply offering people extra time or a little money for getting the right answer does not transform them into political encyclopedias, but it does significantly affect how they answer these questions.

The paper continues as follows. Next, we motivate and explain the experimental design in greater detail. Then, we describe the survey in which the experiment was included. We then present our results, focusing first on how experimental variations affect responses to individual questions and then on how the variations affect the number of questions answered correctly. In

[1] Some survey organizations offer respondents compensation, but it is a reward for completing the survey, not for response quality.

brief discussion and conclusion sections, we spell out further implications for common interpretations of existing political knowledge data and offer ideas for further research.

**Political Knowledge under Realistic Conditions: Experimental Design and Implementation**

To measure political knowledge in surveys, researchers typically use a set of factual questions about politics. When asked to quickly recall a fact, respondents will first draw upon the kind of memory known as "declarative memory." Existing approaches presume that simply asking a political knowledge question will induce respondents to exert sufficient effort to retrieving all relevant facts. Declarative memory, however, does not work in such a fashion. There is a correspondence between the amount of effort one devotes towards recalling facts and the range of facts recalled (National Research Council 1994: 28-29, Kandel, et. al. 1995: 656-664). With minimal effort, a relatively small set of facts from declarative memory will emerge. With greater effort, more facts tend to emerge. One possible limitation of existing political knowledge measures is that they are based on a more limited draw from declarative memory than occurs at key political moments.

A second limitation exists. Cognitive psychologists distinguish fact-based declarative memory from rule-based "procedural memory." Knowing where and how to find things, such as Kerry's vote share in Kansas, is an important form of procedural memory. Procedural memory "accumulates slowly through repetition over many trials, is expressed primarily by improved performance, and cannot ordinarily be expressed in words" (Kandel, et. al. 1995:658).[2] To figure out which candidate we prefer, or how we feel about a new policy proposal, many of us draw on procedural memories of how to gather information that might help us in our decision. The pace

[2] Many scholars (see, e.g., the review in National Research Council 2000) use the terms "declarative" and "procedural" to distinguish the two kinds of memory. Kandel, et. al, (1995:656) refer to declarative memory as "explicit" memory and to procedural memory as "implicit memory."

and incentives that characterize typical survey procedures, in contrast, may inhibit respondents from using such procedural memories.

Typical surveys do not provide a context conducive to documenting the kinds of declarative and procedural knowledge that are relevant to many normatively-driven questions about what citizens know at important democratic moments. To create a survey context that more closely resembles the circumstances of such moments, we experimentally manipulate two elements of the survey interview, the time respondents have to complete a set of political and economic knowledge questions and whether or not they earn money for answering these questions correctly. In so doing, our design integrates elements from leading nationally-representative political surveys, survey experiments, and experimental economics. Figure 1 provides a schematic representation of our design.

[Figure 1 about here.]

The first experimental factor is time. The pace of a typical survey interview prompts respondents to answer questions quickly. To determine the extent to which this attribute affects political knowledge scores, we gave one randomly selected half of our sample only one minute to answer each knowledge question. We gave the other half a broader window—up to 24 hours—to respond. In our Internet-based survey, which respondents completed using a computer or a WebTV unit, the knowledge questions appeared after an initial battery that solicited party identification, interest in politics, and previous turnout. After this battery, all respondents saw a common introduction:

> In the next part of this study, you will be asked 14 questions about politics, public policy, and economics. Many people don't know the answers to these questions, but it is helpful for us if you answer, even if you're not sure what the correct answer is. We encourage you to take a guess on every question. At the end of this study, you will see a summary of how many questions you answered correctly.

Respondents in the "one minute" condition were informed that

> You will have 1 minute to answer each question. After 1 minute, you will be automatically forwarded to the next question. If you finish answering a question before 1 minute is up, you may proceed to the next question by clicking on the 'Next Question' button.

Each of the knowledge questions was programmed to be on screen for up to one minute. If respondents answered the question within that period or if one minute had expired, the screen changed to show the next question. In the "one minute" condition, respondents could not go back to a previous knowledge question after they had moved past it in the interview.

Respondents in the "24 hour" condition were informed that

> You will have 24 hours to answer these questions from the time you see the first question. Once the 24 hours are up or whenever you decide that you are done, you will be forwarded to the next section and will not be able to return to the knowledge questions. However, before you reach the next section, you may go back to previous knowledge questions by clicking the 'back' button.

Starting from the moment at which respondents saw the first knowledge question, they had 24 hours to complete the knowledge series. During this period, they could go back and forth between knowledge questions (but not to the initial questions about interest, turnout, and partisanship), change their answers, and interrupt and resume the survey as often as they liked. When respondents reached the end of the knowledge sequence, a screen informed them that they could modify their answers until their 24 hours were up or move to next part of the survey (at which point they could not return to the knowledge questions.)

The second experimental factor is a monetary reward. To encourage respondents to base their answers on a more extensive search of declarative memory or to draw on relevant aspects of procedural memory, we include in our experimental design an incentive for greater respondent effort. Following a well-established practice in experimental economics, we chose a monetary incentive because

> "The presence and amount of financial incentive does seem to affect average performance in many tasks, particularly judgment tasks where effort responds to incentives…and where increased effort improves performance. Prototypical tasks of this

> sort are memory or recall tasks (in which paying attention helps)…which are so mundane that monetary reward induces persistent diligence when intrinsic motivation wanes." (Camerer and Hogarth 1999:8).

We offered one randomly selected half of our sample a monetary reward, one dollar, for each correct answer.[3] After receiving the timing information (according to their random assignment on the first factor), respondents in the pay conditions received the following instructions:

> In addition, we will pay you for answering questions correctly. You will earn 1,000 bonus points ($1) for every correct answer you give. So, if you answer 3 of the 14 questions correctly, you will earn 3,000 bonus points ($3). If you answer 7 of the 14 questions correctly, you will earn 7,000 bonus points ($7). The more questions you answer correctly, the more you will earn.

The dependent variables in our study come from answers to the 14 knowledge questions that we asked. For open-ended answers, we specified a range of answers we considered close enough to count as correct. Respondents were told the number of questions they answered correctly (and the rewards they had earned) at the very end of the interview. This sequence is necessary because we asked some post-treatment questions about the election and wanted to avoid confounding answers to these questions by offering explicit performance feedback.[4] Respondents received

[3] We chose $1 per question (which amounts to a maximum possible payoff of $14) because we assumed that the amount would be non-trivial for most respondents and because this amount allowed us to stay within our budget while generating a sufficient number of cases per cell for rigorous statistical evaluations. Following Gneezy and Rustichini (1999), we believed that rewards far smaller than $1 per correct answer would not be a sufficient incentive. While rewards greater than $1 per correct answer would likely generate greater treatment effects, they present two problems. First, at some point the incentive is unrealistically large for most respondents. Second, given a fixed budget, higher payments would have forced us to approach a far fewer set of respondents per cell. While many successful experiments in psychology are administered with far fewer respondents per cell than our target (300 per cell) scholarly norms in political science often prompt a demonstration that treatment effects hold for important subgroups and are sustainable in multivariate analyses – both of which we demonstrate later in the paper. With such expectations in mind, we opted for larger cell values and moderately sized payments.

[4] We provided the following explanation to respondents after they had answered the last knowledge question (and, in the "24 hour" condition, decided to move to the next part of the interview): "We are now calculating the number of questions you answered correctly and will show you how you did momentarily.

credit for correct answers in the form of "bonus points." The firm that conducted our study, Knowledge Networks, sends their panelists checks for $25 when they reach 25,000 points (which they can also earn in other surveys they take.) For all practical purposes, we consider our incentives direct cash rewards. The instructions in the pay condition mentioned the bonus points as well as their dollar equivalents. Respondents in the pay conditions were reminded on every screen with a knowledge question that a correct answer would earn them a specific monetary reward.[5]

**The Survey**

Our 2x2 experiment was embedded in a representative survey of U.S. residents conducted by Knowledge Networks between October 19 and November 1, 2004. Knowledge Networks interviews national probability samples over the Internet by providing a large panel, selected through Random Digit Dialing, with WebTV units and/or free Internet connections in exchange for taking surveys. The participants for this study constitute a randomly selected subset of the KN panel and approximate a random sample of the U.S. adult population. Our survey was assigned to 1,550 panelists of whom 1,220 (79 percent) completed it. Eighty percent of the respondents who completed the survey did so within 4 days of the fielding date.[6]

---

In the meantime, please answer a few simple questions that have no right or wrong answers. After you finish the next few questions, we will show you the results."

[5] We conducted a manipulation check to determine if respondents actually spend more time answering the knowledge questions in the "24 hour" conditions. Not surprisingly, giving people 24 hours to complete the interview leads them to take more time. Monetary incentives also increase interview length. This effect kicks in even in the "one minute with pay" condition, but is stronger in the "24 hour" conditions. For details, see the appendix. Later in this paper, we also test for selection effects more explicitly by adding control variables to the analysis.

[6] We examined whether assignment to the experimental conditions affected completion rates (i.e., whether providing extra time for responses or paying respondents for correct answers would affect the likelihood that they complete the entire interview). If it does, then we must estimate this indirect effect of the experimental manipulations as well as their direct effects. Part of this complication is avoided because the assignment of the money factor occurred only when respondents reached the knowledge section of the interview. Respondents who quit the survey before that point could not have been affected by the monetary incentive as we had not revealed that aspect of the survey to them yet. Only seventeen

Knowledge Networks' survey methodology makes our study a conservative test of our hypothesis. The company informs its panelists by email when a new survey is waiting for them. They can then take the survey at a time of their own choosing. Hence, even respondents in our control group ("one minute, no pay") are not literally caught during dinner or at other inopportune moments and asked to answer the knowledge questions on the spot. In fact, they even had the opportunity to pause the interview when they learned that they would be asked political knowledge questions. (However, they could not stop the relevant timers once they saw the first knowledge question.) Clearly, we do not capture the true inconvenience of a typical phone interview. Indirectly, panelists also receive credit for taking a survey because Knowledge Networks pays for their WebTV unit and/or an Internet connection to their PC. To be sure, this reward does not represent an incentive to answer thoughtfully, but the conditions in our control group do not recreate the conditions of a typical phone interview perfectly. Therefore, even respondents in the control group are more motivated and less inconvenienced than respondents in the telephone surveys from which many claims about political knowledge are derived. [7]

---

respondents quit after reaching that point in the interview. Ten were in the "24 hour" condition and may have forgotten to resume the interview with the 24-hour period. Assignment to the time condition was determined at the beginning of the interview but revealed to the respondents only at the beginning of the knowledge sequence. The completion rates in the two time conditions are not statistically different. Eighty percent of the respondents assigned to the "one minute" condition completed the interview, compared to 78 percent in the "24 hour" condition. Of the seventeen respondents who never made it to the knowledge questions, seven would have been assigned to the pay condition and ten to the no pay condition. Hence, selection effects are very unlikely. Therefore, we consider experimental differences between respondents who completed the interview as valid estimates of the true treatment effects.

[7] Survey data usually show learning over the course of a campaign. The percentage of respondents who answer the same knowledge questions correctly increases as the election approaches (eg. Johnston, Hagen, and Jamieson 2004). Because of this learning effect, the timing of our experiment might suppress the effect of our treatments. It was fielded two weeks before Election Day 2004 after a long campaign. If people do poorly on survey-based knowledge tests when there is no immediate real-world reason to know something about politics, then the marginal impact of extra time and monetary incentives would likely be greater if we had not conducted the study in the shadow of an impending election.

The demonstration of learning effects using the standard survey procedure we criticize here raises another question. Is that result alone not enough to demonstrate that people acquire more information in advance of key democratic moments—and that standard phone survey procedure can pick up this learning effect?

The survey included 14 knowledge questions, some of them open-ended, others multiple choice. Seven were about political issues and seven were about economic issues. Our aim was to ask respondents about some of the most focal campaign issues of the 2004 presidential race as well as economic conditions that were relevant in that context. We asked about the candidates' positions on tax cuts, education, and the line-item veto. Our questions covered the Senate vote on the Iraq authorization and the 9/11 commission's findings about links between al-Qaeda and Iraq. We included questions about the composition of the Senate and the competitiveness of the presidential campaign in the respondent's state. Finally, we wanted to know if respondents could tell us how many Americans are not covered by health insurance, how many live in poverty, and how many are unemployed. We tested their knowledge of the estate tax and the federal debt. While we included two questions that we thought would be relatively easy for everyone to answer, we expected most to be difficult. We also included at least one question each about Bush (his proposal to increase funding for the striving readers initiative) and Kerry (his position on the line-item veto) for which the correct answer went against prevailing stereotypes of them. In short, we asked challenging questions about important issues facing voters in the 2004 election. A complete list of questions and their wording is in table 1.

(Table 1 goes here)

We followed Mondak and Davis's (2001) recommendation to discourage "Don't Know" responses by not giving respondents an explicit "Don't Know" option for each question.

---

We believe that standard survey methods fall short for two related reasons. First, this argument applies only political decisions with a fixed time horizon, such as elections or referenda. Yet people also form opinions about new issues, policy proposals, and politicians. Without the focal point of an impending election, it is hard to know when learning occurs and how much information underlies people's newly formed opinions. Second, even in an election context, we do not know when respondents reached their decision. On Election Day, some early deciders may already have forgotten some of the information that affected their decision. Knowledge in standard survey interviews close to Election Day will therefore underestimate what some respondents knew when they made up their minds (Lodge, Steenbergen, and Braun 1995).

Respondents could of course hit the "next question" button without marking any answer, but their responses show that very few of them did (see below). Discouraging "Don't Know" responses reduces distortions because some people are more likely to guess than others in the absence of encouragement. More importantly, our interest in analyzing how variations in the survey context affect the nature and distribution of erroneous responses makes us very interested in how people guess. In other words, answers to open-ended (and some multiple-choice) knowledge questions tell us more than whether people get the answer right or wrong. We can also see if people's answers are off by a lot or a little. Believing that 15 percent of all Americans live in poverty may have different implications for people's decision-making than believing that 35 percent do. For this reason, we start our analysis by looking at the experimental effects on the response distributions for individual questions. Later, we will use a summary knowledge measure for each respondent. For simplicity, our summary measure is the number of correct responses. This choice raises the question of how to determine the range of answers to open-ended questions that we consider correct. These ranges are listed in table 1. We have also run the analyses with different ranges of the same general magnitude and the treatment effects remain similar.[8]

[8] A second coding decision pertains to respondents who did not see all of the knowledge questions. This situation arises in the "24 hour" conditions for respondents who reach the 24-hour time limit before completing the whole battery. In particular, some respondents in those conditions started the knowledge section, took a break, and never returned to complete the remaining questions. In our analysis of individual questions, we exclude respondents who never saw a particular question. Doing so in the calculation of the summary measures, however, would inflate the missing data problems and bias our sample because we would be excluding the less motivated respondents who forgot to finish the questionnaire. Hence, we use the total number of correct answers as our dependent variable and code all non-answered questions as incorrect. Only 24 of the respondents who started the knowledge section did not see all 14 knowledge questions, so this coding decision does not affect the substance of our findings.

## Findings

For illustrative purposes, we begin this presentation by illustrating the experimental effects for two individual questions. We use these questions to highlight useful metrics for understanding what citizens do and do not know. Later, we provide analogous results for all of the questions we asked.

One of our knowledge questions was about a proposal that John Kerry made in his 2004 campaign. We asked about the family income level above which Kerry proposed to eliminate tax cuts that George W. Bush had signed into law. We offered respondents five answers: "Over $50,000 a year," "over $100,000 a year," "over $150,000 a year," "over $200,000 a year," and "over $500,000 a year." Figure 2 shows the responses to this question in the four experimental conditions.

(Figure 2 goes here)

The percentage of respondents who select the correct response (over $200,000 a year) rises from 41 percent in the control condition to 55 percent in the "24 hours with pay" condition. Our two treatments combined thus increases the proportion of respondents who answer the question correctly by 24 percent. Individually, the two treatments also improve performance. A logit analyses shows that each factors has an independent effect on political knowledge (time: $\beta = .21$, $p = .08$; pay: $\beta = .36$, $p = .002$; N=1197). The interaction of time and pay, on the other hand, is insignificant. In other words, the monetary incentive has the same effect regardless of whether respondents has one minute or up to 24 hours to answer the question and vice versa.

Figure 2 shows a second effect of the experimental treatments that we can use to expand our understanding of what citizens do and do not know about politics. The treatments affect the

nature of incorrect responses.[9] The percentage of respondents who believe or, perhaps, guess that Kerry wants to eliminate the Bush tax cut for all families with incomes over $50,000 per year drops from 19 percent in the control condition to between 10 and 13 percent in the three other conditions. A more thorough test of this dynamic comes from examining how the experimental design affects the magnitude of erroneous responses (i.e., deviations from the correct answer). This is a meaningful measure only when response options are ordered. We measure deviations from the correct answer by calculating the absolute difference between the correct answer and the response. For example, a respondent who selected "over $100,000 a year" deviates from the correct response by $100,000. The average absolute error is the mean deviation across respondents. (We multiply the measure by -1 so that higher values indicate greater proximity to the correct answer.)[10]

For the Kerry tax proposal, both time and pay reduce deviations from the correct answer by significant amounts. In the control group, the average absolute error was $72,000. In the "24 hours with pay" condition, this error is only $52,000, a reduction of 27 percent. Time alone reduces the error by 10 percent, while a monetary incentive alone lowers it by 16 percent. The two independent main effects are significant at $p < .05$ (money) and $p < .10$ (time), while the interaction is not significant.

[9] This kind of result parallels that of Smith and Walker (1993) who found that monetary incentives reduce the number of responses caused by subjects who could be characterized as thoughtless and unmotivated.

[10] Here, we define what deviation means for respondents who saw the question, but did not answer it. Even though few people did this in our experiment, omitting them would bias our results because respondents are least likely to provide an answer in the control condition. We deal with this problem by randomly assigning a response to all such respondents. Each response is equally likely to be assigned, so this procedure assumes that respondents would offer a completely random guess if they were forced to select a response, which seems reasonable for people who did not provide an answer even though they were invited to guess. (To reduce the impact of extreme random draws, we repeated each analysis 300 times and report the average of those 300 calculations.)

The usefulness of deviation from the correct answer as a measure of political knowledge is even more apparent for an important class of open-ended questions (i.e., questions where many different numerical-based responses are possible). Figure 3 plots responses to a question about the percentage of Americans that did not have health insurance in 2003 (for the complete text, see table 1.) If respondents entered a number outside the range of possible answers (in this example, a number less than 0 or greater than 100), an error message alerted them and gave them an opportunity to change their response. Otherwise, any response was recorded and the next question appeared. According to the Census Bureau, 15.6 percent of the population did not have health insurance in 2003.

For analytic purposes, we treat answers within 6 points of the correct answer (9.6 to 21.6 percent) as sufficiently close to be 'correct.' By this definition, 13 percent of the respondents in the control condition offered the correct answer, compared to 24 percent in the "24 hours with pay." A logit model yields significant main effects for both time ($p < .01$) and pay ($p < .01$) on the probability of a correct response.

Turning to an examination of the kinds of errors respondents made, the distributions of responses in figure 3 illustrate that respondents vastly overestimated the percentage of uninsured Americans. However, the experimental treatments move the distribution closer to the correct answer. The average absolute deviation from the correct answer decreases from 26 percent in the control group to 22 percent when respondents are offered extra time and a monetary inventive, a 14-percent reduction in average error magnitude. Regressing the absolute error on the experimental factors indicates a significant effect for time ($p < .10$), but not for money or the interaction term.

(Figure 3 goes here)

Having used one multiple-choice question and one open-ended question to highlight means by which we can measure what citizens do and do not know and the impact of the experimental variations on these measures, we now present comparable analyses for all knowledge questions in Tables 2 and 3. Table 2 lists the percentage of correct responses for each question in each experimental condition. For most items, the percentage is lowest in the control condition and highest in the "24 hours with pay" condition. The right-hand side of table 2 provides significance tests for the effects of the two experimental factors based on logit models predicting correct responses. We list *p*-values for the time × money interaction term when it is significant. When it is not, we reestimate the model with only the two main effects and report those *p*-values.

(Table 2 goes here)

The bottom row in table 2 lists the mean percentage of correct answers across all questions. On average, 33 percent of the respondents answered correctly in the control condition. Time and pay together raise this percentage to 41, a 25-percent increase in performance.

Giving respondents more time improved their performance significantly for 10 of the 14 questions. Monetary incentives alone had a significant impact on correct answers to four questions. For one question (about Bush's striving readers initiative), both factors were necessary to increase the proportion of correct responses by a significant amount. The interaction of the two factors was also significant for the estate tax question. In sum, time had a very consistent and statistically significant effect on respondents' ability to answer knowledge questions correctly. The impact of monetary incentives was also positive, and in some cases quite sizeable, but not as consistently significant.

Table 3 shows the average absolute deviation from the correct answer in each condition for questions where an ordering of responses in terms of error magnitudes is feasible. The right half

of the table reports the *p*-values for the two experimental factors from an OLS regression of deviation from the correct answer. The results largely confirm the impact of the experimental treatments seen in table 2. In all but one case, offering more time leads to a statistically significant reduction in error, while the monetary incentive's effect is significant for the two tax-related questions. Across all questions in table 3, as seen in the table's bottom row, simply paying respondents for correct answers, without increasing the amount of time respondents had to answer questions, reduced the magnitude of their errors by an average of nearly 6 percent. Giving them more time to answer questions without paying them reduced errors by approximately 12 percent on average. Combining the two treatments lowered errors by an average of 14 percent.

(Table 3 goes here)

The only question for which we do not find any experimental effects in either table pertains to the 9/11 Commission's findings about the link between Iraq and al-Qaeda. The percentage of respondents who correctly said that the report found evidence that "a few al-Qaeda individuals visited Iraq or had contact with Iraqi officials" was unaffected by the experimental treatments. The experimental treatment does, however, affect the response distribution. The percentage of respondents who answered that there was no connection at all between Iraq and al-Qaeda increased from 26 percent in the control condition to 37 percent in the "24 hours with pay" condition. Both time and money were individually sufficient to lower the likelihood of the incorrect response that "Iraq was directly involved in carrying out the September 11th attacks." Even with additional time and monetary incentives, it is clear that some respondents failed to note the subtleties of the commission's findings (i.e., respondents were less likely to choose a response option that indicated deep involvement but appeared to confound a claim of "no

involvement" with the true claim of "limited involvement"). We think that valuable future analysis on this topic could focus on how various media outlets covered the Iraq-al-Qaeda connection and whether such coverage fed the confusion.

Examining each question individually offers a detailed impression of the experimental results. We now examine the impact of the experimental treatments on a respondent's total number of correct answers. Figure 4 plots the distribution of that variable in the control and the "24 hours with pay" conditions. Adding time and money shift the distribution to the right bringing it to resemble a normal distribution (which is noteworthy given that we purposely chose questions that are much harder than standard political knowledge questions).

(Figure 4 and Table 4 go here)

Table 4 summarizes the distributions in all four experimental conditions. The experimental treatments reduce by nearly half the share of respondents who gets only one or two questions right and nearly doubles the share who answers eight or more questions correctly. Overall, by offering respondents time and a monetary incentive, we raised their average performance by 24 percent. Analysis of variance confirms that both experimental factors significantly increase knowledge scores (Time: $F[1,1216] = 17.2$, $p < .0001$; Pay: $F[1,1216] = 5.5$, $p < .02$). The interaction of the two factors, on the other hand is not significant (Time × Money: $F[1,1216] = .43$, n.s.), indicating that each factor works independently.

To clarify the substantive significance of this 24 percent increase in political knowledge scores, we compare it to the effect of other politically relevant variables. Table 5 presents two OLS models that regress our 15-point knowledge index on the experimental main effects with (column 2) and without (column 1) controls for common demographic and attitudinal predictors of political knowledge. A comparison of the two columns reveals that the effects of time and pay

change only marginally when the non-experimental controls are added (which simply confirms that the experimental randomization worked.)

(Table 5 goes here)

Mirroring earlier results, the combined effect of the two experimental factors amounts to a performance increase of one correct question per respondent (1.09 questions in the regression without controls and 1.01 questions in the regression with controls). Compared to the impact of the control variables, themselves selected for previous success in predicting political knowledge scores, this is a sizable effect. It is almost twice the size of the gender gap in this study and about the equivalent of a two standard deviation gain in a respondent's income. When compared to even more widely cited predictors of political knowledge, the experimental effects are more impressive. Together, they amount to two thirds of the difference between the most and least interested respondents. (The omitted category in table 5 includes respondents who follow politics "only now and then" or "hardly at all.") Similarly, the joint effect of time and pay is equivalent to three quarters of the difference between a college graduate and a respondent with only a high school education. In other words, for $1 per question, we observed a performance increase on par with the results of three years of tuition payments.

In sum, there are differences between the political knowledge of the control group and the treatment groups that are statistically and substantively significant. They are sufficient to reject the null hypothesis that existing findings about political knowledge are robust to changes in the method of eliciting responses. While our treatments by no means suggest that citizens are encyclopedic in their knowledge of politics, they do suggest that traditional means of eliciting knowledge on surveys are likely biased downwards because they inhibit the use of relevant

procedural memories and do not provide sufficient incentive for respondents to draw upon much of their declarative memory.

Our findings suggest the importance of being careful when drawing broad conclusions about citizens' political abilities from the results of existing political knowledge tests. In addition to previously reported problems with the kinds of questions often used in knowledge assessments (Burton and Blair 1991, Gould 1996, Guthrie 1998, Lupia 2006, Tourangeau, et. al. 2000, Chapter 3.3), our experiment reveals that previous observations are also, in part, driven by attributes of the survey context that are quite unlike the contexts in which citizens make important political decisions. While such surveys do represent performance in environments that resemble surveys, we must be careful when generalizing those findings to dissimilar environments. By not allowing respondents to use the kinds of cognitive tools that they might use in real political circumstances, standard surveys systematically underestimate what citizens know at important political moments.

**Discussion: For Whom Do Time and Money Make the Largest and Smallest Difference?**

Previous studies have documented a correspondence between observed political knowledge and a range of personal and contextual factors. We take a moment to report on how our experimental manipulations affect these correspondences. Such analyses can confer two benefits. First, they provide evidence about the robustness of previous claims about the effects of factors such as education, income, race, gender, and interest in politics on political knowledge. Second, for people who seek to increase political knowledge through various kinds of motivational strategies (e.g., rallies, advertisements, or other forms of mobilization), such analyses may indicate the kinds of citizens who will be most and least responsive to civic education appeals whose success is premised on greater opportunity or motivation improving performance. To

show these effects, we present separate OLS regression models for each experimental group in Table 6. If the coefficients for a particular variable are significantly greater for some groups than others, we can conclude that our treatments change the effect of this variable on political knowledge.

(Table 6 goes here)

Table 6 shows that for most variables, offering greater time or money does not have a statistically significant effect on the relationship between the named factor and political knowledge. There are, however, important exceptions. Despite the relatively small sample size per condition, we find significant differences in the effects of several variables on political knowledge.

Extra time significantly increases performance for respondents who are older and those who are college-educated. These results imply that traditional surveys are particularly likely to underestimate what such citizens know. These are also people who we expect may have broader stores of declarative or procedural knowledge on which they draw in circumstances that are more like important political moments and less like a traditional fast-paced survey.

Also noteworthy is the fact that the experimental treatments actually close the performance gap between respondents who report being very interested in politics and those who report being less interested. We demonstrate this closing gap in two ways. First, notice that political interest has its largest effect in the "one minute with pay" condition. When respondents are motivated by the prospect of a small material reward for answering correctly, but have no opportunity to search for a correct answer, the politically interested among them do particularly well. The difference in performance between the moderately interested and the least interested is large and significant in the "one minute, pay" condition, but not in either of the conditions that grant

respondents more time. A similar pattern holds for the difference between moderate and high political interest. All else constant, respondents with low and moderate interest in politics use the extra time to close the gap in performance.

To make these effects more transparent, figure 5 graphs the predicted values for each of the four conditions. We compare predicted political knowledge for a white male college graduate, with mean age and income, full-time employment, and Internet access. The dotted line illustrates that political interest has its largest impact in the "one minute with pay" condition. The difference between the least and most interested respondents corresponds to two items on the knowledge scale, a very large effect by comparisons to the effects of other variables. For the two "24 hour" conditions, this difference is less than one item. This sizable difference suggests to us that politically interested respondents do well on knowledge tests because they have a lot of political information stored in the part of declarative memory that they can retrieve quickly. When we give people time to draw on their procedural memory, political interest is much less important because with sufficient time even the less interested are able to find answers to political questions. Moreover, the significant effect on performance for this group in the "one minute with pay" condition suggests that traditional political knowledge measures do not document what the most interested people know well – as simply offering them a dollar per correct answer boosted their performance. This finding is evidence that traditional means do not motivate even the most interested citizens to try as hard as they can on political knowledge assessments.

(Figure 5 and Table 7 go here)

The other large effect in table 6 concerns race. In the control condition, we find the same racial difference that many previous studies of political knowledge have found: Minorities do

significantly less well. The effect, in our analysis, is about three quarters of an item. In the two pay conditions, however, the minority effect is dramatically increased. Whites answer almost two more questions correctly than minorities. What happened? Examining the experimental effects separately by race in table 7 reveals that our treatments had less of an effect on the knowledge scores of non-whites. The monetary incentive is particularly ineffective. When non-white subjects have only one minute to answer questions, paying them for correct responses decreases performance, while the effect of adding performance incentives in the 24-hour treatment is virtually zero. We are surprised by this finding. While we have found a number of viable explanations, they are not mutually exclusive, nor are we able to distinguish amongst them empirically.[11] The fact that the differences are so apparent in the multivariate analysis in table 6 suggests that the race difference is not an artifact of income, education, or the other variables listed in that table. We are interested in learning more about this difference in future studies.

## Conclusion

Many observers are concerned about what citizens know about politics. Although many normative questions about the implications of political knowledge remain contested, there is a consensus that reliable measurements are socially valuable. This consensus rests on the premise that an informed population can make decisions that are more beneficial to themselves, their families, and the communities in which they live. Attempts to improve citizens' abilities to perform important political tasks depend on what they know, and social science's ability to help

[11] One that seems particularly plausible follows from Steele and Aronson 1995. In a series of laboratory experiments, Steele and his colleagues showed that "Black participants expecting to take a difficult, ability-diagnostic test showed significantly greater cognitive activation of stereotypes about Blacks, greater cognitive concerns about their ability" (1995: 805) and that such concerns "can further undermine performance by undermining motivation and effort" (1995: 809). It may be that our offer to pay for correct answers raised the stakes for performing well and activated analogous inhibitory mechanisms.

improve their performance depends on how well we can assess if they have the knowledge they have at critical democratic moments.

Our motive is to improve the conceptualization and measurement of socially relevant forms of knowledge. Implicit in this approach is the idea that atypical aspects of the typical survey interview context lead to knowledge measures that are less relevant to questions of democratic competence than commonly presumed. Our critique of existing practice has two prongs. First, we claim that existing measures ignore procedural memory, even though citizens rely on it when making important political decisions. Second, we claim that existing measures make a false presumption about the availability of relevant declarative memories during a typical survey interview. To convert this critique into a constructive basis for improved understanding, we remove typical time restrictions on some respondents and offer a monetary incentive to others. These treatments increase the number of correct responses by 11 to 24 percent on average and also reduce the magnitude of error in erroneous responses. At a minimum, such results clarify how much it would take to induce respondents to approach political knowledge questions in a manner that improves their responses. The answer is "not much." Allowing extra time or introducing a small financial incentive ($1 per correct answer) was sufficient to increase performance significantly.

One may respond that, "Of course people do better if they can ask someone else," but for us that is exactly the point. In many of the circumstances about which political scientists and other observers of politics care most, those who cannot instantly recall a particular fact can take a little more time to think about it, ask someone, or look up the answer. Traditional surveys, while having many virtues, prevent or inhibit exactly the kinds of search activities that are in fact strongly encouraged by people who want others to make more informed decisions.

Another possible objection is that citizens are not very motivated to search their memories at key political moments. For the most realistic assessment of political knowledge, it would be necessary to encourage exactly as much searching as occurs at key political moments. The incentive has to be big enough to compensate for the lack of motivation created by the typical survey context, but not so big as to create unrealistic incentives to get it right. We admit that such a calibration of incentives is by no means straightforward. Citizens vary in the extent to which they are intrinsically motivated to answer survey questions under traditional conditions. A monetary incentive that induces one respondent to devote greater effort to answering knowledge questions may have no effect on another. Clearly, other forms of motivation could be used in the dollar's stead. We intend to investigate this calibration issue further in future work.

Of course, we expect that some citizens could be induced to devote more effort to politics even if we offered substantially more time or money -- but our experiment suggests that performance-related "tipping points" are not difficult to reach for many people. So while low political knowledge scores are sometimes described as if they are a trait, or purely a function of ability, our results undermine such representations. For if political knowledge was purely a function of ability, we should not be able to reject the null hypothesis (as ability should be constant across all conditions by virtue of randomization). In other words, if people do not perform well on knowledge tests because they are stupid and unable to understand politics regardless of how we ask the questions, then the procedural variations we introduced should make no difference. The fact that both experimental factors have sizable effects illustrates that limited ability is not the only obstacle to answering conventional political knowledge questions correctly.

## Appendix: Manipulation Check

We expected that additional time to answer the knowledge questions and a monetary incentive for answering them correctly will lead some respondents to devote greater effort to answering the questions. Reading the question carefully and trying to come up with the right answer might not necessarily take respondents much longer than answering the question "of the top of the head." Only if respondents ask friends or family members or look up the answers, should they take considerably longer than respondents in the control condition or respondents who were given only one minute per question. In short, while we did not necessarily expect huge differences in completion time, both experimental factors should increase interview time beyond that of the control group.

To assess the validity of these expectations, we conducted a manipulation check. Respondents in all conditions could stop and later resume the interview in the parts that precede and follow the knowledge part. That is why even a few respondents in the "one minute" condition took several hours to finish the interview. A measure of interview length is only available for the entire interview, not for the knowledge portion alone. In order to use the total interview time as a meaningful manipulation check, we have to assume that the time respondents spent on the parts before and after the knowledge section is, in expectation, the same in all experimental conditions. We cannot include the 17 respondents who answered knowledge questions, but did not complete the last section of the interview. Knowledge Networks did not collect data on when they interrupted the interview.

Our check reveals that giving people more time to answer the knowledge questions led some, but not all respondents to spend more time completing the survey. The average length of the interview (for interviews under one hour) is 14.6 minutes in the "24 hour no pay" condition and

16.6 minutes in the "24 hour with pay" condition, compared to 11.9 minutes in the control condition. The effect of monetary incentives on interview time is less obvious. Regressing logged interview time on the two factors and their interaction yields a strong main effect for time, and smaller, non-significant effects for pay and the interaction. Depending on how the dependent variable is specified (logged, truncated, or recoded to reduce the influence of very long interview times), the interaction term increases to approach statistical significance. When the interaction term is omitted, the main effect for pay increases and becomes significant. We interpret these results as follows: Not surprisingly, giving people 24 hours to complete the interview leads them to take more time. (This is true even for interviews of less than an hour.) Monetary incentives also increase interview length. This effect kicks in even in the "one minute" condition, but is stronger in the "24 hour" conditions.

**References**


Bennett, Stephen Earl. 1995. Comparing Americans' political information in 1988 and 1992. *Journal of Politics* 57 (2):521-532.

Blair, Edward A., and Scot Burton. 1987. "Cognitive Processes Used By Survey Respondents to Answer Behavioral Frequency Questions." *Journal of Consumer Research* 14: 280-288.

Burton, Scot, and Edward Blair. 1991. "Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys." *Public Opinion Quarterly* 55: 50-79.

Camerer, Colin F. and Robin M. Hogarth. 1999. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty* 19: 7-42.

Converse, Philip E. 1964. The nature of belief systems in mass publics. In *Ideology and Discontent*, edited by D. E. Apter. New York: Free Press.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know About Politics and Why it Matters*. New Haven: Yale University Press.

Gould, Steven Jay. 1996. *The Mismeasure of Man*. New York: Norton.

Gneezy, Uri, and Aldo Rustichini. 1999. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115: 79-810.

Guthrie, Robert V. 1998. *Even the Rat Was White: A Historical View of Psychology*. Boston; Allyn and Bacon.

Hertwig, Ralph, and Andreas Ortmann. 2001. Experimental Practices in Economics: A Methodological Challenge for Psychologists?" *Behavioral and Brain Sciences* 24: 383-451.

Johnston, Richard, Michael Hagen, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics.* New York: Cambridge University Press.

Kandel, Eric R., James H. Schwartz, and Thomas M. Jessell. 1995. *Essentials of Neural Science and Behavior*. Norwalk, CT: Appleton and Lange.

Kennickell, Arthur B. 2000. "Asymmetric Information, Interviewer Behavior, and Unit Nonresponse." Working Paper, The Federal Reserve Board (Washington DC).

Kennickell, Arthur B. 2003. "Reordering the Darkness: Application of Effort and Unit Nonresponse in the Survey of Consumer Finances." Working Paper, The Federal Reserve Board (Washington DC).

Kinder, Donald R., and David O. Sears. 1985. Public opinion and political action. In *The Handbook of social psychology* (pp. 659-741), 3rd edition, edited by G. Lindzey and E. Aronson. New York: Random House.

Krosnick, Jon A., and D. Alwin. 1987. "An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement." *Public Opinion Quarterly* 51: 201-219.

Lodge, Milton, M. Steenbergen, and S. Braun. 1995. The responsive voter: Campaign information and the dynamics or candidate evaluation. *American Political Science Review* 89:309-326.

Lupia, Arthur. 2006. "How Elitism Undermines the Study of Political Knowledge." Forthcoming, *Critical Review*.

Luskin, Robert C. 1987. "Measuring Political Sophistication." *American Journal of Political Science* 31: 856-899.

Mondak, Jeffery J., and Belinda Creel Davis. 2001. "Asked and Answered: Knowledge Levels When We Won't Take 'Don't Know' for an Answer." 2001. *Political Behavior* 23: 199-224.

National Research Council. 1994. "Transfer: Training for Performance." In Druckman, Daniel, and Robert A. Bjork, eds. *Learning, Remembering, Believing: Enhancing Human Performance*. Washington, DC: National Academy Press.

Schwarz, Norbert. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Smith, Vernon L., and James N. Walker. 1993. "Monetary Rewards and Decision Cost in Experimental Economics." *Economic Inquiry* 31: 245-261.

Steele, Claude, and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African-Americans." *Journal of Personality and Social Psychology* 69: 797-811.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of the Survey Response*. New York: Cambridge University Press.

Table 1: Knowledge Questions

| Question ID | Question wording | Response options (Correct response in bold) |
|---|---|---|
| Senate term | How long is the term of office for a U.S. Senator? | **open-ended, correct: 6** |
| Reps in Senate | Of the 100 members of the U.S. Senate, how many are members of the Republican party? | **open-ended, correct: 51, accepted range: 51-59**. This range reflects two key points: a Republican majority and its inability to prevent filibusters. |
| Closeness in 2000 | What was the outcome of the 2000 Presidential Election in the state in which you now live?<br><br>[The correct answer depends on the respondent's residence.] | • Bush won by more than 5 percentage points<br>• Bush won by less than 5 percentage points<br>• Gore won by less than 5 percentage points<br>• Gore won by more than 5 percentage points |
| Striving Readers | President Bush proposed a "Striving Readers initiative" to help high school students who are not reading as well as they should be for their age. What is the status of the Striving Readers program? | • The program was implemented in 2002 and has already led to a 1.3% increase in functional literacy among high school students.<br>• **President Bush has proposed to fund this program at $100 million in his 2005 budget.**<br>• President Bush proposed this program, but did not include any funding for it in his 2005 budget.<br>• The program started last year, but in his 2005 budget President Bush proposed to cut its funding by $200 million. |
| Iraq authorization | In the key Senate vote on October 11, 2002, how many Democratic Senators voted to give President Bush the authority to attack Iraq? | • None of them<br>• Two Democratic senators<br>• About a quarter of all Democratic senators<br>• **A majority of all Democrats in the Senate, but not all of them**<br>• All Democratic senators |
| Line-item veto | A line-item veto allows the president to sign a budget bill while cutting specific spending items and tax expenditures that he disapproves. The Supreme Court recently ruled one version of the line-item veto unconstitutional. Other versions of the line-item veto are less likely to be overruled by the court. Which of the following statements best describes the presidential candidates' positions on new versions of the line item veto? | • President Bush and Senator Kerry both oppose the line-item veto.<br>• President Bush supports a line-item veto, while Senator Kerry opposes it.<br>• Senator Kerry supports a line-item veto, while President Bush opposes it.<br>• **President Bush and Senator Kerry both support a line-item veto.** |

(Table continued on next page)

Table 1: Knowledge Questions (cont.)

| | | |
|---|---|---|
| Al-Qaeda connection | As you may know, a special government commission—called the "9/11 Commission," investigated the circumstances surrounding the September 11 attacks and recently issued its final report. Which statement most accurately represents the Commission's conclusions about the relationship between Iraq and al Qaeda? | • They had no connection at all<br>• **A few al-Qaeda individuals visited Iraq or had contact with Iraqi officials**<br>• Iraq gave substantial financial support to al-Qaeda, but was not involved in the September 11th attacks<br>• Iraq was directly involved in carrying out the September 11th attacks |
| Taxes compared to Europe | Compared with the citizens of Western European countries, do you think Americans pay a higher percentage of their income in taxes, a smaller percentage of their income in taxes, or about the same percentage of their income in taxes? | • A higher percentage<br>• **A smaller percentage**<br>• About the same percentage |
| Unemployment rate | The U.S. Bureau of Labor Statistics counts a person as unemployed if they are not employed at any job and are looking for work. By this definition, what percentage of Americans was unemployed in August of 2004? | • around 11 percent<br>• around 9 percent<br>• around 7 percent<br>• **around 5 percent**<br>• around 3 percent |
| Estate tax | There is a federal estate tax – that is, a tax on the money people leave to others when they die. What percentage of Americans leaves enough money to others for the federal estate tax to kick in? | • About 95% of all Americans<br>• About 70% of all Americans<br>• About 50% of all Americans<br>• About 25% of all Americans<br>• **Less than 5% of all Americans** |
| Uninsured Americans | In August 2004, the United States Census Bureau reported an estimate of the number of Americans without health insurance. The Census Bureau classified people as uninsured if they were not covered by any type of health insurance at any time in 2003. By this definition, what percentage of Americans did not have health insurance in 2003? | **open-ended, correct: 15.6, accepted range: ±6 points** |

(Table continued on next page)

Table 1: Knowledge Questions (cont.)

| | | |
|---|---|---|
| Federal Debt | The outstanding public debt of the United States is the total amount of money owed by the federal government. Every year the government runs a deficit, the size of the public debt grows. Every year the government runs a surplus, the size of the public debt shrinks. In January of 2001, when President Bush took office, the outstanding public debt of the United States was approximately 5.7 trillion dollars. Which of the following responses is closest to the outstanding public debt today? | • Less than 3.5 trillion dollars<br>• 4.5 trillion dollars<br>• 5.5 trillion dollars<br>• 6.5 trillion dollars<br>• **7.5 trillion dollars**<br>• 8.5 trillion dollars<br>• More than 9.5 trillion dollars |
| Kerry tax proposal | John Kerry says that he would eliminate the Bush tax cuts on families making how much money? | • Over 50,000 a year<br>• Over 100,000 a year<br>• Over 150,000 a year<br>• **Over 200,000 a year**<br>• Over 500,000 a year |
| Poverty rate | In August 2004, the Census Bureau reported how many Americans live in poverty. The poverty threshold depends on the size of the household. For example, a person under age 65 is considered to live in poverty if his or her 2003 income was below $9,573 and a family of four is considered to live in poverty if its 2003 income was below $18,810. By this definition, what percentage of Americans lived in poverty in 2003? | **open-ended, correct 12.5, accepted range: ±6 points** |

Table 2: Experimental Effects, by Question

| | Percent Correct | | | | Statistical Significance |
|---|---|---|---|---|---|
| | One Minute | | 24 Hours | | |
| | *No Pay* | *Pay* | *No Pay* | *Pay* | |
| Senate term (OE) | 33 | 35 | 39 | 37 | |
| Reps in Senate (OE) | 39 | 45 | 49 | 51 | Time: $p<.01$ |
| Closeness in 2000 | 46 | 54 | 50 | 60 | Time: $p<.10$, Pay: $p<.01$ |
| Striving Readers | 15 | 12 | 17 | 21 | Time × Pay: $p<.10$ |
| Iraq authorization | 56 | 58 | 63 | 62 | Time: $p<.10$ |
| Line-item veto | 39 | 35 | 45 | 44 | Time: $p<.05$ |
| Al-Qaeda connection | 30 | 29 | 30 | 28 | |
| Taxes compared to Europe | 41 | 43 | 52 | 52 | Time: $p<.01$ |
| Unemployment rate | 27 | 34 | 38 | 39 | Time: $p<.01$ |
| Estate tax | 36 | 47 | 46 | 46 | Time: $p<.01$, Pay: $p<.01$, Time × Pay: $p<.05$ |
| Uninsured Americans (OE) | 13 | 16 | 17 | 24 | Time: $p<.01$, Pay: $p<.01$ |
| Federal Debt | 22 | 24 | 27 | 26 | |
| Kerry tax proposal | 41 | 50 | 46 | 55 | Time: $p<.10$, Pay: $p<.01$ |
| Poverty rate (OE) | 18 | 23 | 26 | 28 | Time: $p<.05$ |
| ***Mean Percent Correct*** | ***32.6*** | ***36.1*** | ***38.9*** | ***40.9*** | |

Note: Significance tests are based on logit models regressing response on the two experimental factors and their interaction. If the interaction term in the initial logit model was not significant, the model is reestimated with only the two main effects. (OE) = Open-ended

Table 3: Experimental Effects on Average Absolute Error, by Question

| | **Average Absolute Error** | **Change in Average Absolute Error** | | | **Statistical Significance** |
|---|---|---|---|---|---|
| | One Minute *No Pay* | One Minute *Pay* | 24 Hours *No Pay* | 24 Hours *Pay* | |
| Senate term (OE) | 1.73 | - 0% | - 8.7% | - 7.5% | Time: $p<.05$ |
| Reps in Senate (OE) | 7.67 | + 4.4% | - 19.9% | - 15.5% | Time: $p<.05$ |
| Unemployment rate | 2.71 | - 9.2% | - 15.9% | - 13.3% | Time: $p<.05$ |
| Estate tax | 23.0 | - 23% | - 18.3% | - 20.4% | Time: $p<.05$, Pay: $p<.05$, Time × Pay: $p<.10$ |
| Uninsured Americans (OE) | 25.9 | + 1.9% | - 5.8% | - 13.5% | Time: $p<.05$ |
| Federal Debt | 1.55 | + 0.6% | - 9.7% | - 6.5% | Time: $p<.10$ |
| Kerry tax proposal | 71.5 | - 16.1% | - 10.3% | - 27.4% | Time: $p<.10$, Pay: $p<.05$ |
| Poverty rate (OE) | 21.3 | - 3.8% | - 3.3% | - 9.4% | |
| ***Mean Change in Error*** | | ***- 5.7%*** | ***- 11.5%*** | ***- 14.2%*** | |

Note: Average Absolute Error is the mean absolute deviation from the correct answer. If a respondent did not provide any answer, the response is imputed based on a uniform probability over the entire range of possible responses. We explain this procedure in the text. Significance tests are based on OLS models regressing respondents' absolute deviation from the correct answer on the two experimental factors and their interaction. If the interaction term in the initial logit model was not significant, the model is reestimated with only the two main effects.

Table 4: Experimental Effects, Knowledge Scales

| | One Minute | | 24 Hours | |
|---|---|---|---|---|
| | No Pay | Pay | No Pay | Pay |
| Number of Correct Answers | | | | |
| *less than 2* | 28 % | 21 % | 18 % | 15 % |
| *more than 8* | 10 % | 15 % | 16 % | 19 % |
| Mean Correct | 4.5 | 5.0 | 5.4 | 5.6 |
| **Percent Increase over Control** | | **+ 11 %** | **+ 18 %** | **+ 24 %** |
| Standard Deviation | 2.78 | 2.95 | 2.95 | 3.10 |

Table 5: Comparing Experimental Effects to Other Correlates of Political Knowledge

| | Political Knowledge Score (0-14) | |
|---|---|---|
| Pay Condition: $1 per correct answer | .50*<br>(.24) | .53**<br>(.20) |
| Time Condition: 24 hours | .81***<br>(.24) | .78***<br>(.20) |
| Pay X Time | -.22<br>(.34) | -.30<br>(.29) |
| Follows politics "some of the time" | | .55**<br>(.19) |
| Follows politics "most of the time" | | 1.55***<br>(.20) |
| High school degree only | | .50*<br>(.22) |
| Some college | | 1.19***<br>(.24) |
| College or graduate degree | | 1.83***<br>(.24) |
| Female | | -.57***<br>(.15) |
| Age | | .013**<br>(.005) |
| Racial/Ethnic Minority | | -1.39***<br>(.18) |
| Income (1-19) | | .12***<br>(.02) |
| Full-time employment | | -.35*<br>(.17) |
| Internet access (other than WebTV) | | .52**<br>(.15) |
| Constant | 4.54***<br>(.17) | 1.54***<br>(.41) |
| $R^2$ | .02 | .29 |
| N | 1220 | 1220 |

* $p < .05$, ** $p < .01$

Table 6: Comparing Experimental Effects to Other Correlates of Political Knowledge

| | One Minute | | 24 Hours | |
|---|---|---|---|---|
| | No Pay | Pay | No Pay | Pay |
| Follows politics "some of the time" | -.01 (.34) [a] | 1.43*** (.36) [ab] | .76 (.40) | .05 (.41) [b] |
| Follows politics "most of the time" | 1.78*** (.37) | 2.24*** (.39) [ab] | 1.20** (.43) [a] | 1.01* (.44) [b] |
| High school degree only | .16 (.38) | .54 (.43) | .64 (.47) | .44 (.50) |
| Some college | .42 (.45) [a] | .90* (.45) [b] | .84 (.51) [c] | 2.32*** (.56) [abc] |
| College or graduate degree | 1.69*** (.42) | 1.62** (.46) | 1.83*** (.49) | 2.14*** (.56) |
| Female | -.41 (.27) | -.90** (.28) | -.53 (.32) | -.40 (.32) |
| Age | -.008 (.009) [ac] | .001 (.009) [b] | .040** (.01) [ab] | .023* (.01) [c] |
| Racial/Ethnic Minority | -.76* (.33) [ab] | -1.83*** (.34) [a] | -1.06** (.37) [c] | -1.96*** (.40) [bc] |
| Income (1-19) | .17*** (.04) | .11** (.04) | .098* (.045) | .12** (.05) |
| Full-time employment | -.69* (.31) | -.48 (.30) | -.15 (.38) | -.22 (.35) |
| Internet access (other than WebTV) | .40 (.28) | .51 (.29) | .38 (.33) | .41 (.34) |
| Constant | 2.53*** (.72) | 2.69*** (.70) | 1.41 (.93) | 2.19* (.85) |
| $R^2$ | .36 | .38 | .23 | .28 |
| N | 310 | 304 | 294 | 298 |

*** $p < .001$, ** $p < .01$, * $p < .05$

Note: For comparisons between columns, coefficients with common superscript letters are statistically different from each other at $p < .10$.

Table 7: Experimental Effects for Whites and Non-Whites, Knowledge Scales

| | One Minute | | 24 Hours | |
|---|---|---|---|---|
| | No Pay | Pay | No Pay | Pay |
| *White Respondents (N=941)* | | | | |
| Mean Correct | 4.7 | 5.6 | 5.7 | 6.0 |
| Percent Increase over Control | | **+ 17 %** | **+ 21 %** | **+ 27 %** |
| *Other Respondents (N=279)* | | | | |
| Mean Correct | 3.9 | 3.3 | 4.2 | 4.2 |
| Percent Increase over Control | | **- 13 %** | **+ 9 %** | **+ 8 %** |

Figure 1: The Experiment

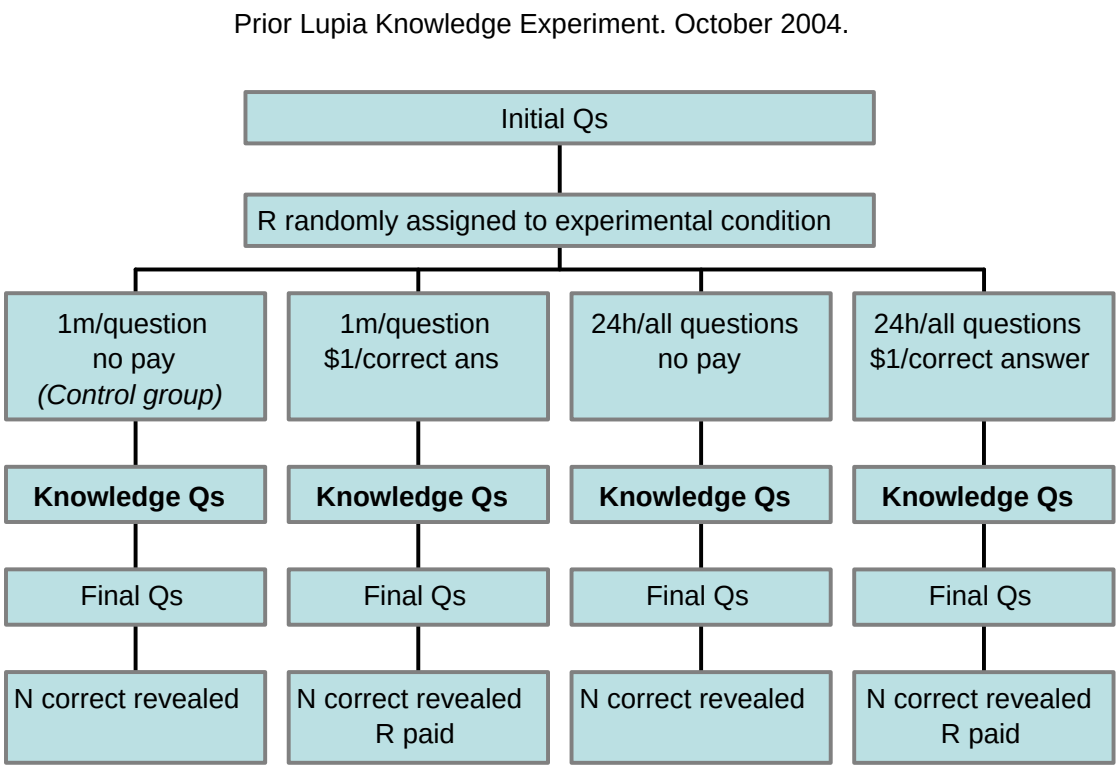


Figure 2: Knowledge of Kerry's Tax Proposal

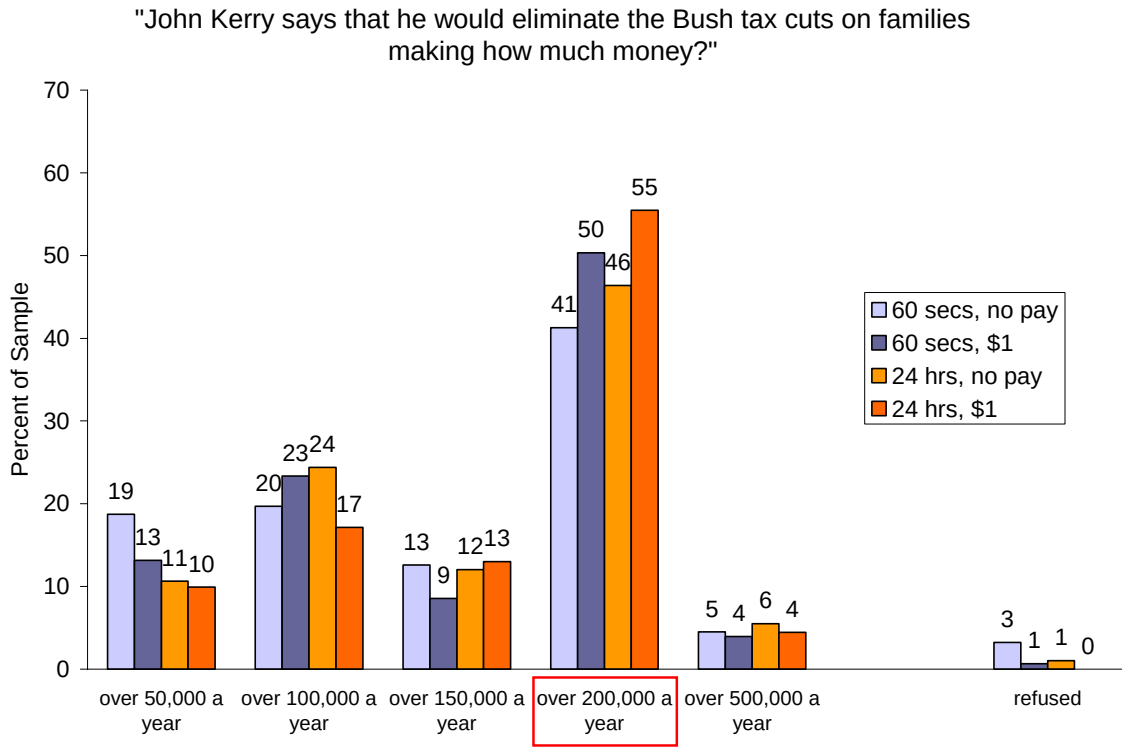
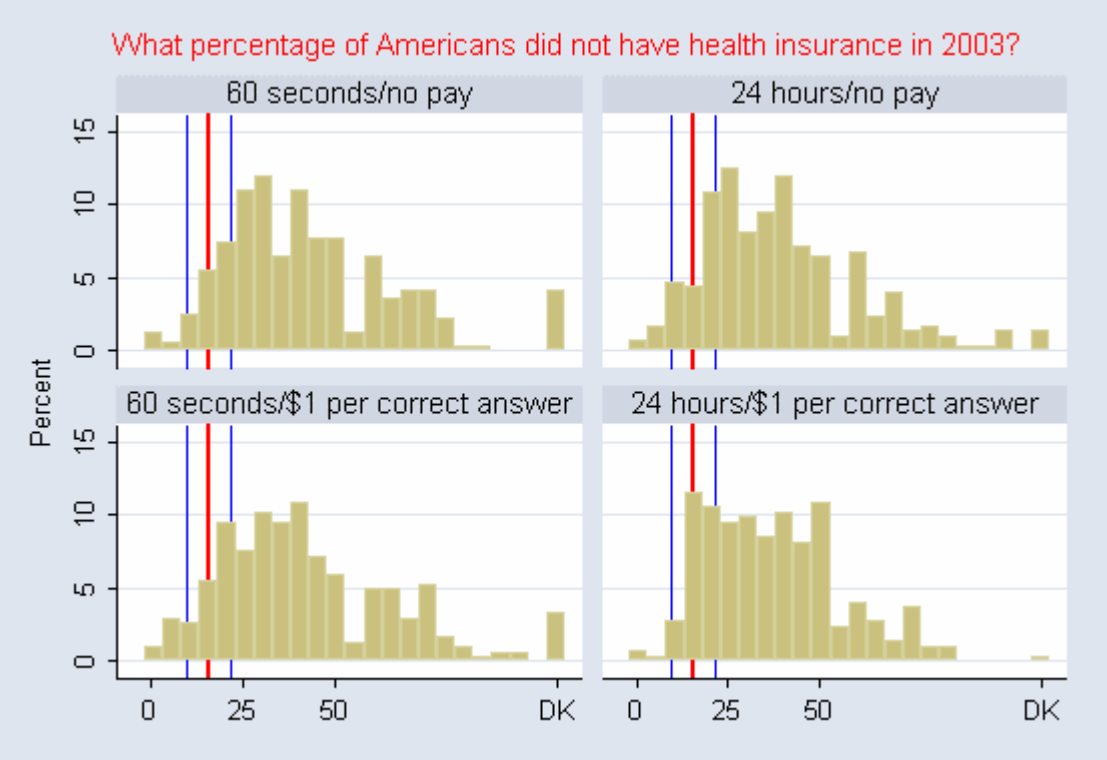


Figure 3: Knowledge of Health Insurance Coverage




Note: The red line marks the correct answer (15.6 percent). For analytic purposes, we accepted answers within the two blues lines (+/- 6 percent) as correct.

Figure 4: Number of Correct Responses, by Experimental Condition

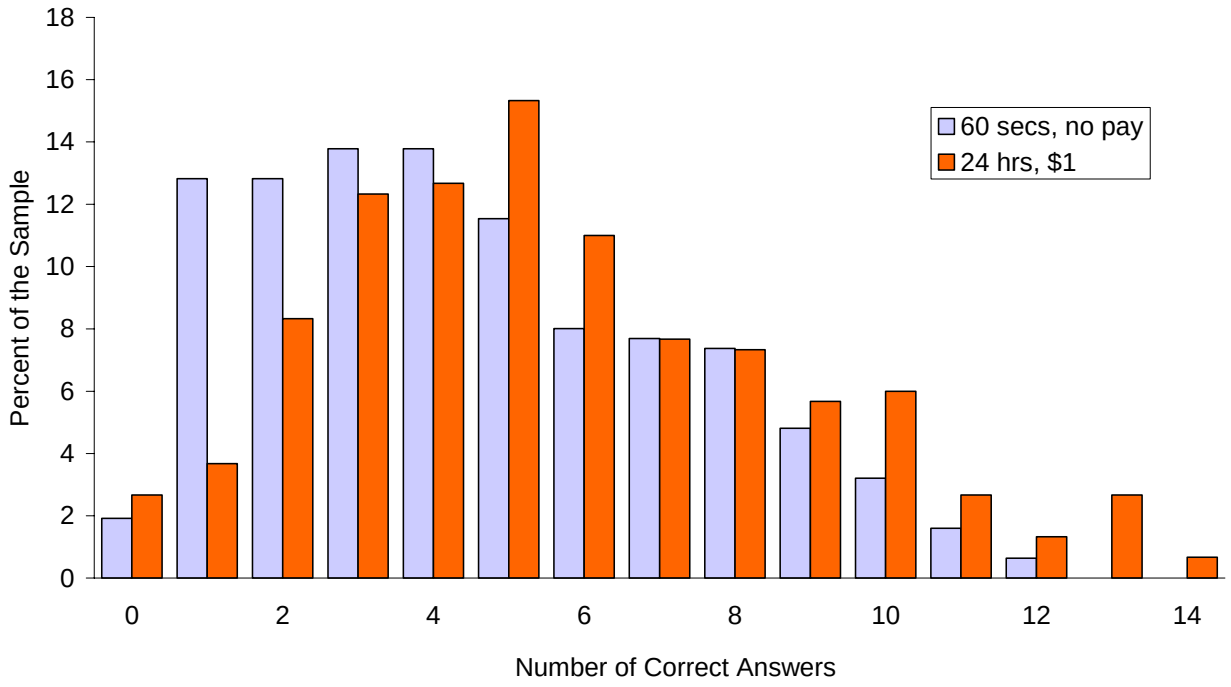


Figure 5: Predicted Political Knowledge, by Condition and Political Interest