

Arthur Lupia  
University of Michigan

# Transparency and Credibility in Election Surveys

# YOUR CHALLENGE

- Goal: Credible scientific contributions
- Constraint: limited time, pressure to publish
- Temptation: Ignore your own assumptions

# My Argument

- Every survey-based claim is based on important assumptions
- Limited introspection about your assumptions puts you at risk of making **easily falsifiable claims**
- Understanding your assumptions, and making them public can increase **your credibility**

# Outline

1. Transparency in Analysis
2. Transparency in Production
3. Example: What Voters Know

The Problem is More than Misspecification

# Analytic Transparency

# Example 1. "Paradigm"

- Use a commercial statistical package to analyze survey data.
- Assume that the underlying causal model is some variant of  $Y = a + b'X + e$ ,
  - $Y = a + b_1(\text{race}) + b_2(\text{race} * \text{something}) + b_3(\text{something}) + b_4 \dots m(\text{other things}) + e$

**What are you assuming about  
this process?**



## Example 2. “Stargazing”

- Begin with a hunch that a particular variable has an unappreciated association with voting behavior or turnout.
- Run a regression. Look for “stars.”
- If the stars support the hunch, STOP
- Otherwise, run more regressions.
- Stop “when the stars align.”

No easily defensible theory guides these decisions.

**What are you assuming about  
this process?**



# Example 3. One Cause?

- Claim: " $X_{-1}$  has no effect on  $Y$ ."
- Evidence:  $X_{-1}$ 's coefficient is not "significant."
  - So,  $X_{-1}$  does not have a statistically significant effect *within the stated model*.
- What if the true underlying data generating mechanism is not identical to the structure of the stated model?

**What are you assuming about  
this process?**



# Remedies

- Existing:
  - new estimators,
  - greater rigor,
  - replication data
- Alternative
  - Track your assumptions by adding a “lab book” norm.

# “Lab Book” Typical Entries

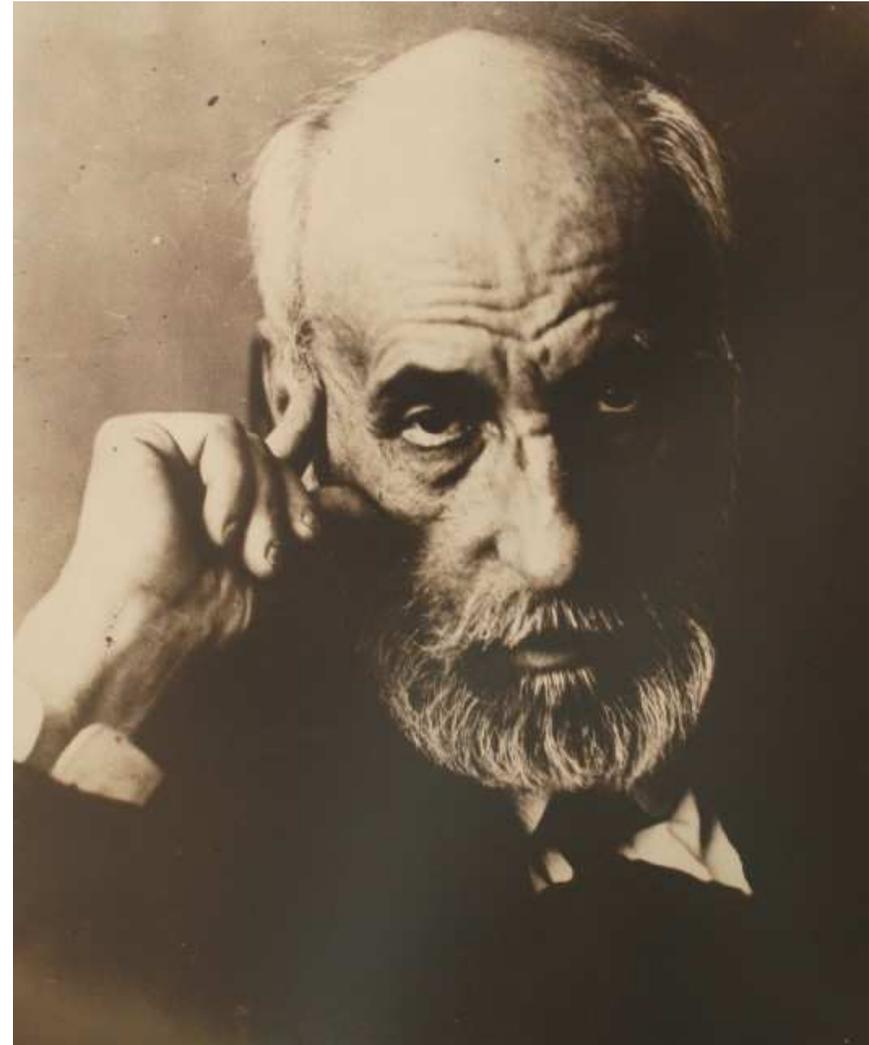
- State the theory
- Explain how specific hypotheses are derived from the theory
- Explain the criteria by which data for evaluating the focal hypotheses were selected or created.
- State the empirical model to be used for the evaluation *in advance*
  - Offer an explicit defense of why a given structure and set of control variables is included.

# “Lab Book” Result Reporting

- Report the results of the initial estimation.
- If the finding causes a change to the theory, data, or model,
  - explain why the changes were *necessary or sufficient* to generate a more reliable inference.
- Do this for every subsequent observation.

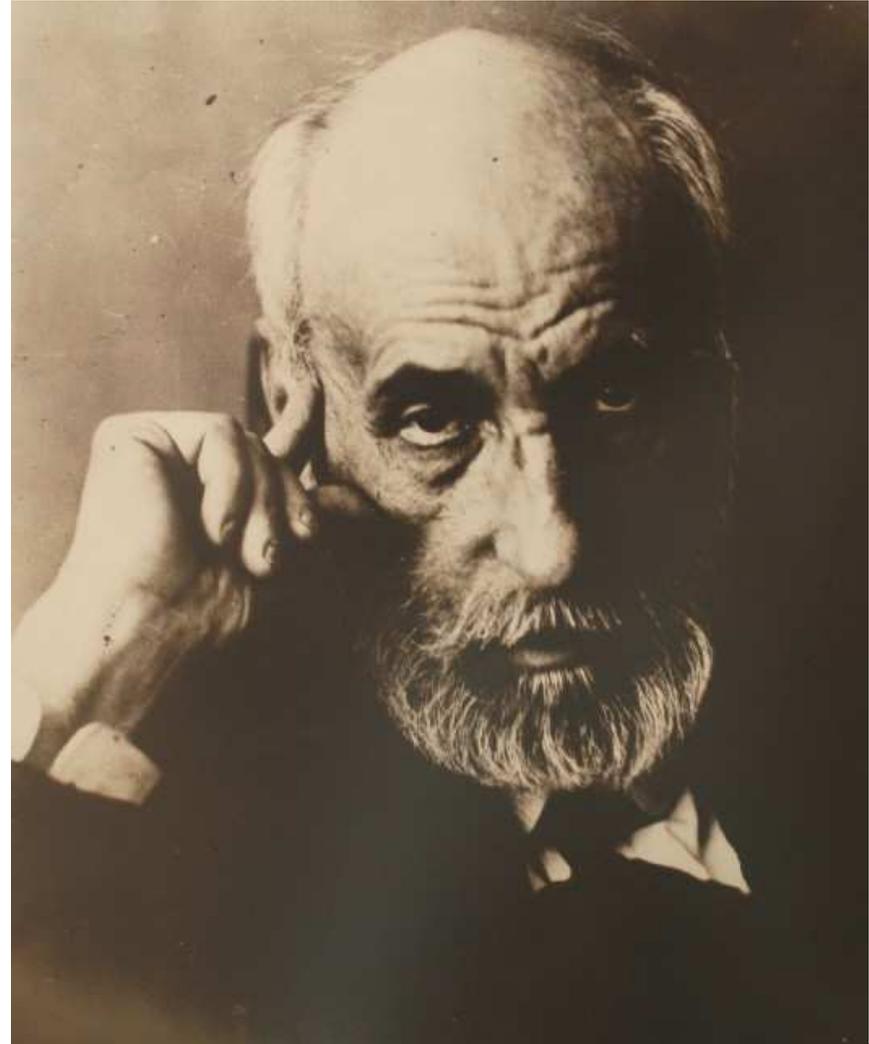
# Santiago Ramon y Cajal (1916)

- “What a wonderful stimulant it would be for the beginner if his instructor, instead of amazing and dismaying him with the sublimity of great past achievements, would reveal instead the origin of each scientific discovery...”



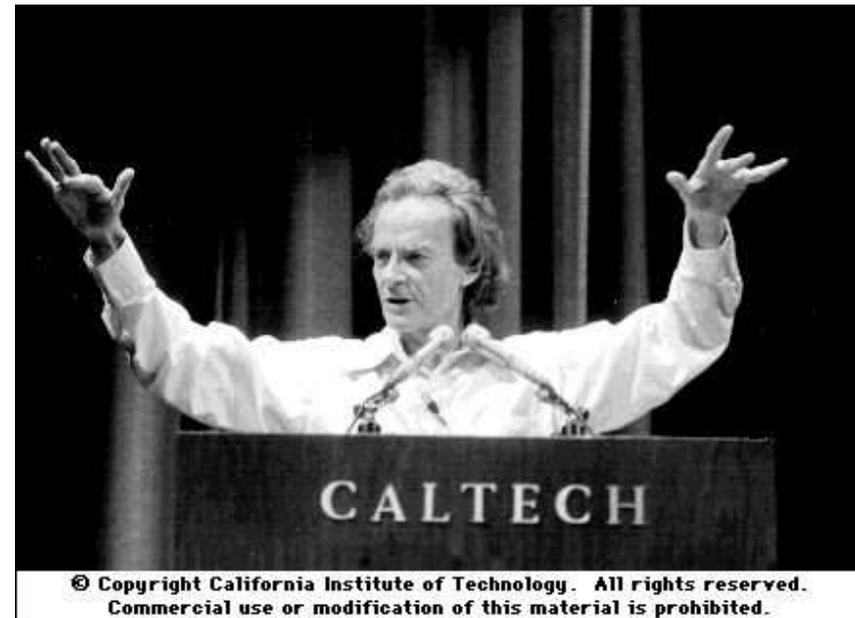
# Santiago Ramon y Cajal (1916)

- “What a wonderful stimulant it would be for the beginner if his instructor, instead of amazing and dismaying him with the sublimity of great past achievements, would reveal instead **the origin of each scientific discovery,**
- the series of errors and missteps that preceded it – information that, from a human perspective, is **essential to an accurate explanation of the discovery.**”



# Richard Feynman (1974 – Caltech Commencement Address)

- [Scientific integrity] corresponds to a kind of **utter honesty**—a kind of leaning over backwards....
- In summary, the idea is **to give all of the information** to help others judge the value of your contribution; not just the information that leads to judgment in one particular direction...



TESS and the ANES Online Commons

# Production Transparency

# Opaqueness in Production

- Can you draw credible inferences from survey data?
- In every survey, a complex decision path converts capital into datapoints.
- Typically, many elements of this path are not public.

# Production “Lab Books”

- Producers start with an ideal
- They have to (or choose to) adjust
- Best practices: “publish” procedures
  - But not how they made decisions
- If “lab books” were also provided, users can draw more accurate inferences

# The Open Source Method of Technological Development

- “The basic idea behind open source is very simple...
- People **improve** it, people **adapt** it, people **fix** bugs.
- And [progress] can happen at a speed that... seems astonishing.”



- Goal: help scholars conduct innovative experiments
- Successful proposals are run on large, national samples *at no charge* to the proposer.
- Any faculty member or graduate student in any social science department anywhere in the world is eligible to make a proposal.
- Proposals may come from any social scientific area as they utilize experimental or quasi-experimental designs.
- TESS has run over 200 experiments for scholars



The Online Commons is a new way to develop election surveys.

- Scholars can
  - *post proposals, comment* on existing proposals, and
  - *amend* their previously posted proposals.
  
- The OC provides broad and public advice about question salience and credibility.
  
- From 2006-2008, over 600 scholars proposed over 3000 questions.

from

*The Trouble with Voters...and Those Who Try to Fix Them*

# **Improving Political Knowledge Assessments**

# Political Knowledge Measures

- Widely cited in the popular and academic presses.
- Many validity and interpretative problems.
- Three sources of subjectivity
  - Question selection
  - Determination of which answers are correct
  - Interpretation of data

# Advice

- Design:
  - What is the task?
  - Is the fact necessary or sufficient for completing the task?
- Analysis:
  - If you want to claim that knowledge or ignorance causes a behavior – what is the counterfactual?

Which questions should we “put on the test?”

# What must a voter know?

Arthur Lupia  
University of Michigan

# Breakdown: Technical

- Competence is with respect to a task
  - Example: make a binary choice.
  - Example: Write a constitution.
- A voter is competent if:
  - She makes the same choice she would have made given different (e.g., more) information.

# How to Build a Competence Measure

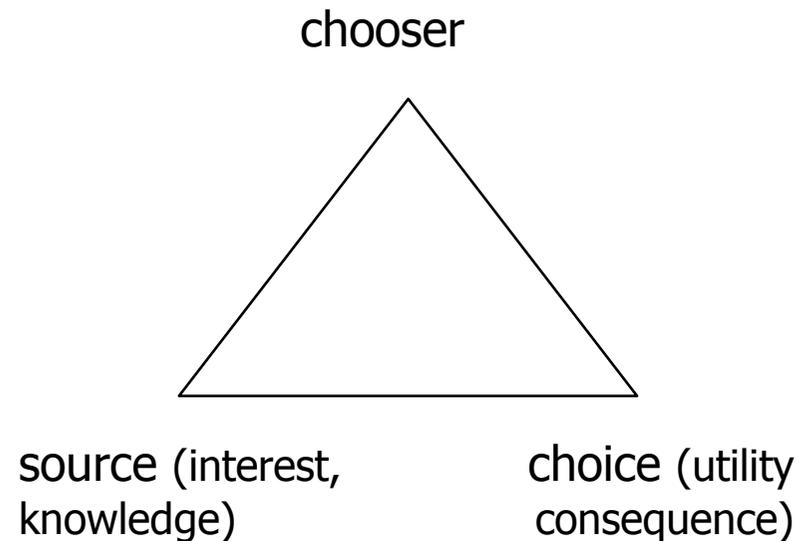
- What is the task?
- Is the fact *necessary* or *sufficient* for completing the task?

# Of Necessity...

- The most informative questions are those for which knowledge is necessary for accomplishing the task,  $t$ .
- Proxies can remove “necessary” status from focal claims.
- As a fact diverges from “necessary,” knowledge of it is a weaker competence indicator

# Sufficient & Efficient

- A proxy is sufficient for competence only in the presence of other information.
- Proxies are efficient when they are effective and require a *relatively small* investment.



# Example of Sufficiency

- A voter makes choices in 20 two-candidate elections.
  - She can cast 1,048,576 possible ballots.
- If “competence” requires the “correct” vote *in all 20 cases*, is any proxy sufficient?

The American Civic Literacy Program

# **Subjectivity in Question Selection**

# The American Civic Literacy Program

- “Students Know Less After 4 College Years.”
- “Critics have long expressed their concern that the nation’s colleges are not teaching what students need to know to *effectively participate in the American political process.*” – ISI Press Release
- *“This is useful knowledge we are talking about.”*  
Josiah Bunting – ISI Chairman

# Sample ISI Questions

- 1. Jamestown, Virginia, was first settled by Europeans during which period?
  - a) 1301–1400
  - b) 1401–1500
  - c) 1501–1600
  - d) 1601–1700
  - e) 1701–1800
- 26. The Declaration of Independence relies most obviously on the political thought of:
  - a) Plato.
  - b) Niccolo Machiavelli.
  - c) David Hume.
  - d) John Locke.
  - e) Georg Hegel.

The IEA Civic Education Study

# **Subjectivity in Determining Which Answers Are Correct**

# IEA Civic Education Study

- “Study of young teens finds civics ranks low; 28-nation survey bodes ill for citizenship” *Washington Times*
- “The assessment items were developed to measure knowledge and understanding of key principles that are universal across democracies.” IEA
- “US students scored significantly above the national average on the total civic knowledge scale. Furthermore, in no country did students significantly outperform US students.” IEA

## SOURCE: IEA Civic Education Study, Standard Population of 14-Year-Olds Tested in 1999

The item shown in Figure 2.10 is an example of a skills item based on a political cartoon. It asks students to identify the message of the cartoon about history books. Seventy-nine percent of U.S. students answered this item correctly. The international average on this item was 57 percent. Hence, although this was a relatively hard item internationally, it was much easier for U.S. students.

Figure 2.10.—Example skills Item 2



What is the message or main point of this cartoon? History textbooks...

- are sometimes changed to avoid mentioning problematic events from the past.
- for children must be shorter than books written for adults.
- are full of information that is not interesting.
- should be written using a computer and not a pencil.

Correct answer: A U.S. average: 79% International average: 57%

SOURCE: IEA Civic Education Study, Standard Population of 14-Year-Olds Tested in 1999.

## **SOURCE: IEA Civic Education Study, Standard Population of 14-Year-Olds Tested in 1999**

**A woman who has a young child is interviewed for a job at a travel agency. Which of the following is an example of discrimination? She does not get the job because...**

- A** She has no previous experience.
- B** She is a mother.
- C** She speaks only one language.
- D** She demands a high salary.

*Correct answer: B*

*U.S. average: 80%. International average: 50%.*

# Are They Valid?

Factor-analytic validity tests do not identify this problem, see, e.g., Gould (1996: 280) in *The Mismeasure of Man*

- The first principal component is a mathematical abstraction that can be calculated for any matrix of correlation coefficients; it is not a "thing" with physical reality.
- ... such a claim can never arise from the mathematics alone, only from additional knowledge of the physical nature of the measures themselves.

# What We Should Know

- "the selection of specific items remains fairly subjective, guided by the goals of the research and influenced by factors not easily quantified" (Delli Carpini and Keeter 1996: 299)
- Greater procedural transparency is needed

# ANES Political Knowledge Questions

- How selected?
- How validated?
- Which answers are correct?

A Lupia, AS Levine, JO Menning, G Sin  
*Perspectives on Politics 2007*

**Were Bush Tax Cut Supporters  
“Simply Ignorant?” A Second Look at Conservatives  
and Liberals in “Homer Gets a Tax Cut”**

# “Homer Gets a Tax Cut”

- “[T]he strong plurality support for Bush’s tax cut...is entirely attributable to simple ignorance.”
  - Bartels (2005:24)

# The data: both papers

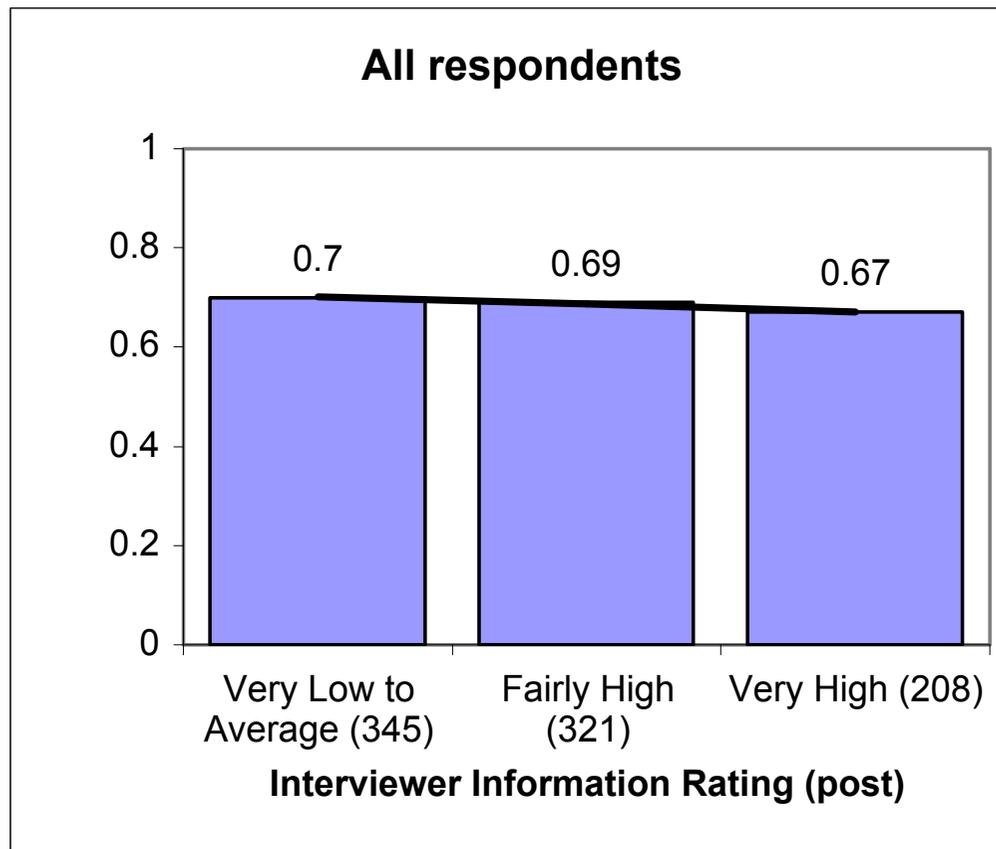
- Key question
  - “As you may recall, Congress passed [President Bush signed] a big tax cut last year. Did you favor or oppose the tax cut, or is this something you haven’t thought about?”

# Homer: Multivariate Analysis

Variable	"Homer"
Political Information (0 to 1)	<b>-.907</b> (.353)
Republican Party ID (-1 to 1)	.759 (.055)
Family Income (in 1000s)	.0002 (.001)
"President Bush" wording	-.080 (.049)
Constant	.873 (.153)
Obs	896

Dependent variable: Support for/Opposition to the Tax Cut

# One View of Information's Impact

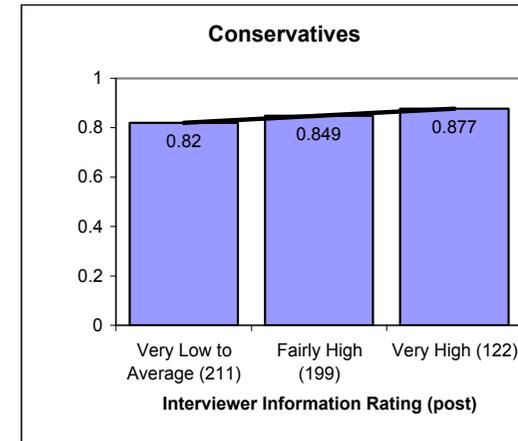
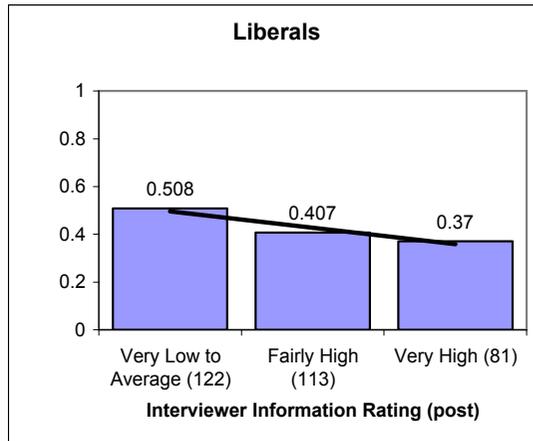


**Dependent variable: Support for/Opposition to the Tax Cut**

# Homer: Key Assumption

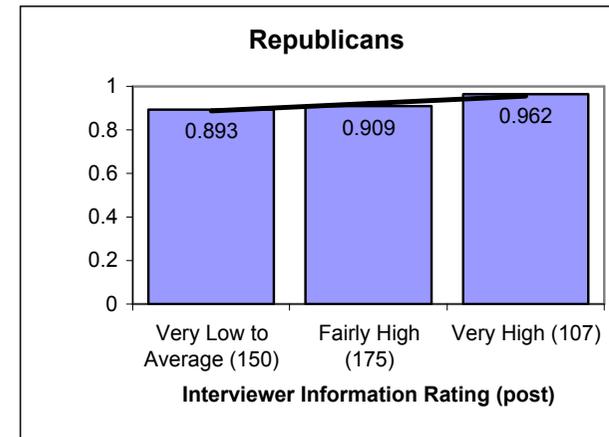
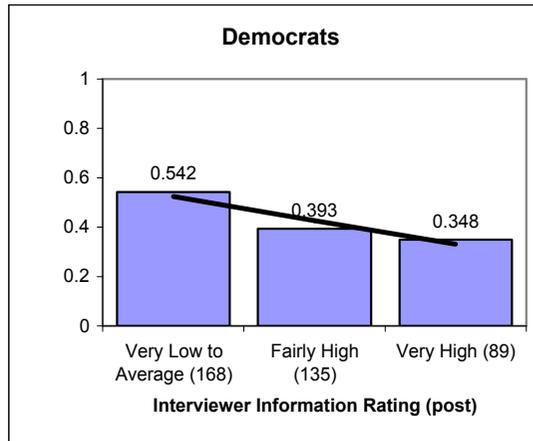
- Everyone -- whether liberal or conservative, Republican or Democrat – responds to more information level in the same way.
- Question:
  - Is it possible that on issues such as the merit of a tax cut proposal, reasonable people can disagree?

# Another View...



- Most conservatives, and over one-third of the most informed liberals, supported the tax cut

# And Another



- Over 96% of the most informed Republicans supported the tax cut

# Conclusion

- The “Homer” result is due to:
  - Presuming that people with different worldviews react to “more information” in the same way.
  - Liberal opinions on this issue vary far more than those of other respondents.
  
- The explanation need not be “simple ignorance.”

Markus Prior and Arthur Lupia  
*Am. J. Polit. Sci.*, Jan 08

**Money, Time,  
and the Ability to Answer Political  
Knowledge Questions**

# Types of memory

- Declarative
  - When asked to recall a fact in surveys, respondents draw upon a kind of memory called “declarative memory.”
  - The ability to draw declarative memory is a function of motivation.
- Procedural
  - Knowing how and where to find things.

# Proposition

- Survey-based knowledge tests catch respondents by surprise.
  - They capture “quick recall”
  
- “Quick recall” ≠ “Political learning skills”

# Research Design

- A two-by-two design.
- Factor 1: Money
  - Control: **no incentive**
  - Treatment: **\$1** per correct answer
- Factor 2: Time
  - Control: **60 seconds** per question
  - Treatment: **24 hours** to complete all questions

# Research Design

- Embedded in national survey
  - conducted by Knowledge Networks
  - Oct. 19 – Nov. 1, 2004
  - N = 1,550 (1,220 completes, 79%)
- The questions
  - 7 on political issues, 7 on economic issues
  - relevant to the 2004 election
  - Most are “hard”

Table 1: Effect of a Monetary Incentive on the Number of Correct Responses to Knowledge

Questions

	No Pay	Pay	Percent Increase
Mean Correct	4.5	5.0*	<b>+ 11</b>
Standard Deviation	2.78	2.95	
N	312	306	
<b>Follows politics ...</b>			
“most of the time” (N=205)	6.2	6.5	+ 4
“some of the time” (N=222)	3.9	5.2**	+ 32
“only now and then” or “hardly at all” (N=189)	3.5	3.3	- 4
<b>College Degree (N=182)</b>			
No College Degree (N=436)	6.1	6.5	+ 7
	3.9	4.5*	+ 15
<b>Female (N=321)</b>			
Male (N=297)	4.1	4.5	+ 8
	4.9	5.8*	+ 17
<b>Age</b>			
18 - 34 (N=150)	4.6	4.4	- 4
35 - 59 (N=291)	4.5	5.4*	+ 20
60 - (N=177)	4.6	5.1	+ 10
<b>White (N=477)</b>			
Non-whites (N=141)	4.7	5.6**	+ 17
	3.9	3.3	- 13
<b>Works full time (N=341)</b>			
Does not work full time (N=277)	4.5	5.1	+ 11
	4.5	5.0	+ 11
<b>Married (N=371)</b>			
Not married (N=247)	4.8	5.3	+ 9
	4.1	4.7	+ 15

\*  $p < .05$ , \*\*  $p < .01$  (two-tailed t-test)

Note: All respondents had one minute to complete each knowledge question. For significant experimental effects, the percent increase is bolded.

5.6

5.4

5.2

5

4.8

4.6

4.4

4.2

4

**Control**

**Pay, 60s**

**No pay,  
24 hrs**

**Pay, 24  
hrs**

+11%

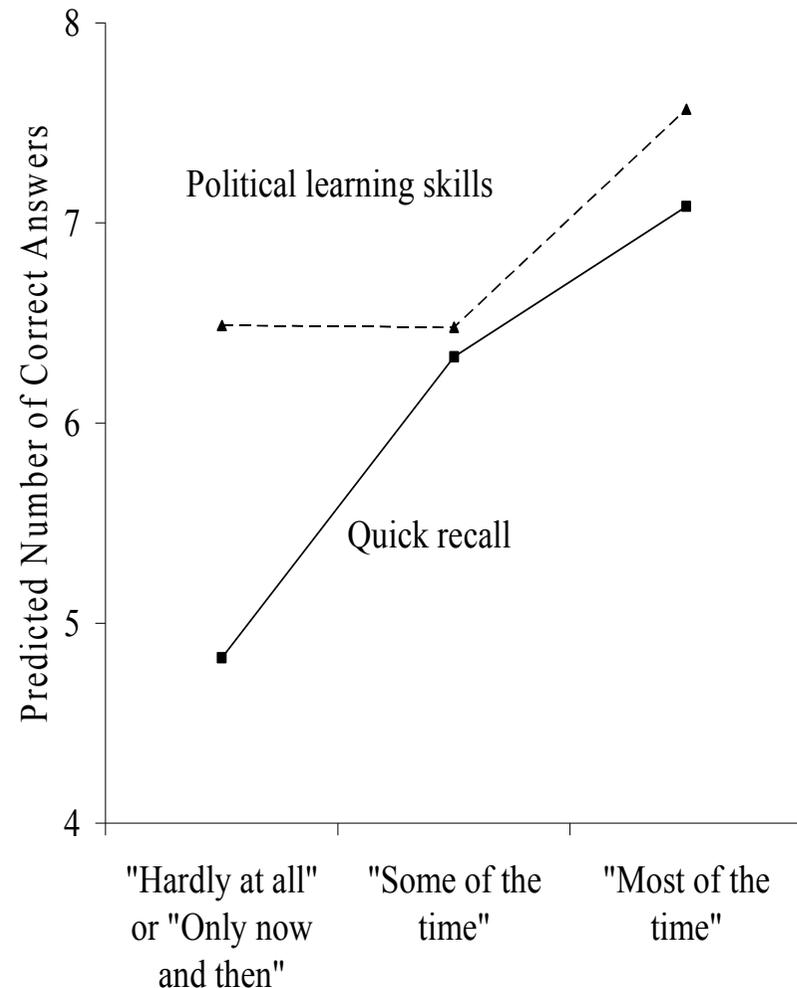
+18%

+24%

-  **Control**
-  **Pay, 60s**
-  **No pay, 24 hrs**
-  **Pay, 24 hrs**

# Closing the Gap

- Paying respondents closes the knowledge gap between highly and moderately interested citizens.
- This result suggests that moderately interested citizens know more than they tell us in surveys.



**"How often would you say you follow what's going on in government and public affairs?"**

# Advice – Political Knowledge

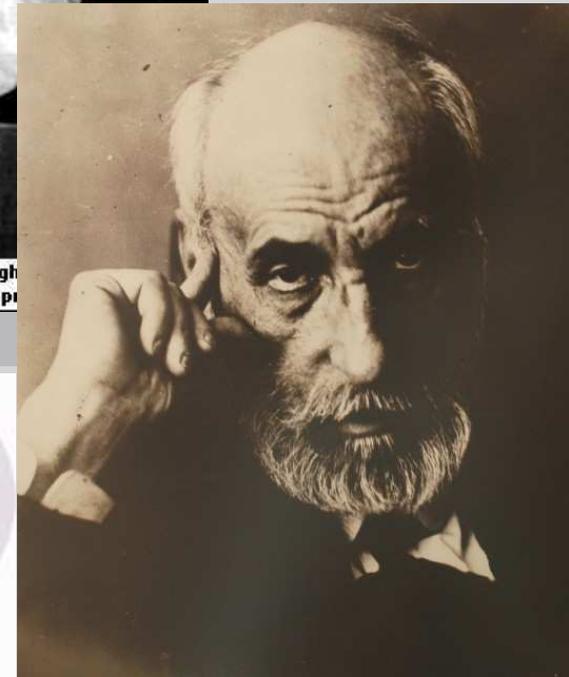
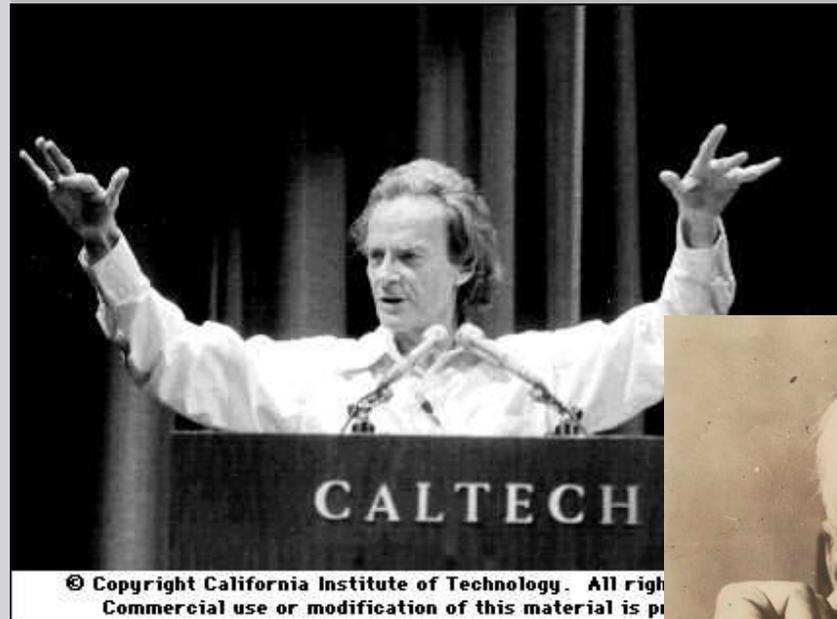
- Design:
  - What is the task?
  - Is the fact necessary or sufficient for completing the task?
- Analysis:
  - If you want to claim that knowledge or ignorance causes a behavior – what is the counterfactual?
- Result: Greater credibility for your research

## My Argument

Every survey-based claim is based on important assumptions.

Know your assumptions, make your procedures public

Transparency is a key to your credibility



**Thank you!**