www.sciencemag.org/cgi/content/full/331/6016/404/DC1

# Supporting Online Material for

## Defeating Creationism in the Courtroom, But Not in the Classroom

Michael B. Berkman and Eric Plutzer*

*To whom correspondence should be addressed. E-mail: plutzer@psu.edu

Published 28 January 2011, *Science* **331**, 404 (2011)
DOI: 10.1126/science.1198902

**This PDF file includes**



**Other Supporting Online Material for this manuscript includes the following:**
(available at www.sciencemag.org/cgi/content/full/331/6016/404/DC1)

A zipped archive containing the following::

*National Survey of High School Teachers: Codebook, Questionnaire, and Data File* (*.pdf)
Questionnaire Data File as an Excel document (*.xlsx) and a data (*.dta) file
Statistical Analysis System formats and export files (*.xpf and *.xpt)

# Supporting Online Material

## Materials and Methods

### Sampling and Data Collection Methods

The data for this paper come from a survey of U.S. high school biology teachers conducted between March 5 and May 1, 2007. The data collection procedures were approved by the Pennsylvania State University Institutional Review Board (IRB #24387). More precisely, the data are from two simultaneous studies using identical questionnaires and overlapping sampling frames. One study was a mail-only study with teachers selected randomly from a database maintained by Quality Education Data, Inc. The database contains names and school mailing addresses for more than 80% of public school teachers in the United States. To be eligible for selection, teachers needed to be in a public school that included grades 9 and 10 and had to include "biology," "life sciences," or "AP biology" among their 1–6 job descriptors in the database. This operational definition permitted the surveying of educators whose primary field was in another science or in science support (e.g., computer lab coordinator); this group of teachers made up 2% of those invited to participate. Completed surveys were discarded if the educator did not teach a high school biology class during the previous year or had recently retired. Following the *Tailored Design Method* for mail surveys (*S1*), five hundred teachers received a prenotification letter, a survey packet (with a two dollar bill and postage paid return envelope), a reminder postcard, and a replacement packet. We received 200 completed questionnaires for a return rate of 40%.

The second study is based on 1500 names drawn from a *subset* of the original database based on the availability of a working email address. This allowed us to not only include all features from the first study but also two additional email follow-up reminders. These emails included a link to a Web version of the survey, making this survey "multimodal" (*S1*, *S2*). A total of 739 respondents completed the multimodal study for a return rate of 49%. Respondents from both surveys are combined in all analyses reported in this paper. After excluding 58 "out-of-scope" respondents (e.g., bad address, no longer teaching, not a biology teacher), the return rate for the combined data set is 48% (926/1942).

**The response rate** is calculated using the formula RR4 in (*S3*):

$$RR4 = \frac{(Completes)}{(Completes) + (Refusals) + e(Unknown)}$$

Where:

**RR4** = American Association for Public Opinion Research (AAPOR) response rate #4

**Completes** = Completed survey returned or Web survey completed

**Refusals** = Blank questionnaires returned without explanation of ineligibility, retirement, or bad address

**Unknown** = Unknown disposition of mailings that were not returned

**e** = Estimated proportion of cases of unknown eligibility that are eligible

Based on third-party audits of the reliability of commercially available mailing lists of teachers and based on ineligible respondents who returned questionnaires, we estimate that 5% of the 1011 unreturned questionnaires were undeliverable or received by teachers ineligible for the survey (setting *e=0.95*). Assuming that all remaining unreturned mailings represent eligible respondents, we estimate the overall response rate as 50%.

Response rate in context

Response rates for most kinds of surveys in nearly every academic discipline have declined in the last two decades (*S4*, *S5*, *S6*). The same is true for surveys of teachers. Gallup's *Phi Delta Kappa Survey of Teachers* is not a scholarly research project but is nevertheless illustrative of mail surveys of teachers conducted by the best survey research organizations. The 1986 survey mailed a questionnaire to 2000 teachers, sent a single reminder postcard, and achieved a response rate of 42% (*S7*). The *same methodology* produced response rates of 26% in 1996, and 18% in 1999 (*S8*). Carefully executed academic surveys with extensive follow-ups (like ours) have yielded response rates somewhat higher. Hess, Maranto and Millman's 1995 mail survey of California teachers, for example, yielded a response rate of 42% (*S9*).

However, the survey research field has moved away from response rates as the primary indicator of validity and unbiasedness (*S10*). Indeed, a number of empirical assessments have demonstrated that parameter bias does not necessarily increase as response rates decline (*S11*, *S12*). Rather, it is preferred to make direct estimates of response bias and of how well the sample reproduces known population parameters for key social and demographic variables (*S10*, *S11*). As we note below, the 926 valid, returned questionnaires are representative of the population of inference, and results can be generalized to that population with confidence.

Representativeness of the Sample

The combined data set contains teachers from 49 states (no teachers from Wyoming) and the District of Columbia. We examined the combined sample to determine if it was representative of the population of inference. Although there is no census of teachers, the Common Core of Data (CCD) compiled by the U.S. Department of Education's National Center for Educational Statistics is a census of schools that we used to see if the schools that our sample members taught in were representative of high schools nationally. Our teachers taught at schools with an average enrollment of 1311 students, compared with the national average of 1332. Similarly, our sample's schools had an average of 26% of students qualifying for free lunch (a common indicator of the number of students from poor families) as compared with 28% nationally. Our sample underrepresents teachers at schools with more than 10% Black students (27% compared with 42% nationally) and more than 10% Hispanic students (22% versus 39%). In addition, we slightly overrepresent Midwestern (35% versus 24%) and small town schools (11% versus 8%).

Poststratification weighting

Following standard practice for adjusting surveys for differential nonresponse rates across groups, we calculated poststratification weights to adjust the survey for these (*S10*, *S13*, *S14*). Almost all major government and academic surveys and media polls use survey weights so that analyses match known population distributions. Population surveys may use census data, political-party registration files, or government statistics to serve as a baseline distribution for key demographic or geographic variables. If a group is underrepresented in the final sample of respondents, for example, that group's weight will be greater than one. In particular, weights should be proportional to the inverse of the probability of selection, although with several variables (e.g., region, size of community, racial distribution, and economic status) weights cannot compensate for all sources of nonresponse. However, the reported results using weighted

and unweighted results never differ by more than 3%, and the same substantive conclusions emerge with either weighted or unweighted analyses. This is seen in Table S1, which reports the summary statistics used to construct Figure 1 using both weighted and unweighted data.

**Table S1. Self reports of qualifications of teachers classified by their approach to teaching evolution**

| Sample | Data quality | Advocate of creationism (%) | Advocate of neither (%) | Advocate of evolutionary biology (%) |
|---|---|---|---|---|
| Percentage of the entire sample | (raw data) | 13 | 60 | 28 |
| | (weighted) | 13 | 61 | 26 |
| Percentage within each group who completed a course on evolution | (raw data) | 33 | 37 | 56 |
| | (weighted) | 36 | 38 | 59 |
| Percentage within each group who rate themselves exceptional | (raw data) | 10 | 7 | 26 |
| | (weighted) | 11 | 9 | 25 |

Because the results are similar and because poststratification weights are sample dependent, unweighted data are employed for all analyses in this paper.

Additional concerns about differential nonresponse

Although we find no evidence of differential nonresponse by school characteristics such as the percentage receiving free lunch, we cannot rule out the possibility of differential nonresponse within schools. For example, it is possible that *conditional on school characteristics* advocates for creationism responded at a 40% rate, and advocates for evolutionary biology responded at a 60% rate within the *same kind* of school (e.g, within the group of predominantly black, urban schools); this would create biased descriptive statistics that would not be corrected by poststratification weighting. This kind of bias is theoretically possible in all sample surveys. However, several teachers took the time to tell us that they could not tell if our survey was sponsored by supporters of evolution or of creationism or intelligent design, which we take as a sign that respondents could not discern an agenda that might lead them to decline to participate for ideological or value-based reasons.

The Questionnaire

Teachers completed a survey containing questions about the content of their most recently taught biology course, more specific questions about the teaching of evolution in particular, and a variety of background questions. The actual questionnaire is provided as a separate file, and was designed in the context of previous studies. We adapted question wording from other studies whenever possible and appropriate [a thorough review of this literature can be found in (*S14*)]. The final pencil and paper questionnaire was six pages long and took about 15 minutes to complete; the Web version was identical in content

Although most questions had a "forced choice," several permitted teachers to give open-ended answers. These were transcribed verbatim, and all quotations appearing in the commentary are unedited. To guaranty the confidentiality of responses, quoted respondents are only identified by the state in which they teach.

## SOM Text

## Supporting Information for Figure 1

**Figure 1** is based on three classification variables. The major classification, Approach to Teaching Evolution," is based on answers to six questions.

*Advocates of creationism* had to meet two criteria: (1) They had to report devoting at least one hour to creationism or intelligent design, in answer to the following question: *"Thinking about how you lay out your Biology course for the year, please indicate how many class hours you typically spend on each of the following areas."* One of these areas was *"Intelligent design or creationism."* (2) Teachers had to respond *"agree"* or *"strongly agree"* to at least one of the following questions: *"When I do teach about creationism or intelligent design* (*including answering student questions*), *I emphasize that this is a valid, scientific alternative to Darwinian explanations for the origin of species."* Or *"When I do teach creationism or intelligent design* (*including answering student questions*), *I emphasize that many reputable scientists view these as valid alternatives to Darwinian theory."*

*Advocates of evolutionary biology* had to adopt the pro-evolution stance on all of the following three questions [by responding "agree" for questions (a) and (b), "disagree" for

question (c)], and adopt the strong pro-evolution stance on at least two [by responding "strongly agree" for questions (a) and/or (b), and/or "strongly disagree" for question (c)]: (a) *"When I do teach evolution* (*including answering student questions*)*, I emphasize the broad consensus that evolution is fact even as scientists disagree about the specific mechanisms through which evolution occurred."* (b) *"Evolution serves as the unifying theme for the content of the course."* (c) *"I believe it is possible to offer an excellent general biology course for high school students that includes no mention of Darwin or evolutionary theory."*

The middle row labeled "*Advocate of neither*" is composed of all other teachers who responded to the survey, including all those who left one of the criterion questions blank.

The self-rating question reported by the blue bars was, *"I would rate my knowledge of the scientific evidence bearing on the validity of evolutionary theory as:* (*a*) *Exceptional, on par with many college-level instructors,* (*b*) *Very good compared to most high school biology teachers,* (*c*) *Typical of most high school biology teachers,* (*d*) *I know less about this topic than many other high school biology teachers."* The figure reports the percentage of all teachers who gave the first response.

The percentages depicted by the red bars are based on answers to *"Have you had a specific college-level course in evolution?"*

## In-Text Reports of Statistical Analyses

In the text, we also report a simple test of association between the teacher classification and reports of nervousness. The latter was obtained on the basis of answers to the following question: "*In preparing this questionnaire, we talked to a large number of teachers and other science educators. They have told us about many strategies that teachers use in teaching their high school biology classes and they shared many experiences they have had. Some of these appear below. Please tell us whether you have ever done the following:"* One of the options was, *"I have been nervous about an open house event or a meeting with parents because I believed that I would receive complaints about the teaching of evolution.*"

**Replication Data**

The following additional materials are available to permit interested scholars to replicate and extend the analyses reported here and in our published work:

1. Codebook/Questionnaire documentation. This PDF file contains a brief description of the study design, descriptions of each variable in the public use data set, and the actual pencil and paper questionnaire mailed to all sampled teachers.

2. Replication data set containing all 84 forced-choice responses from the survey. This is available as a Stata data set (*.dta format) and as an Excel file (*.xlsx format). In addition to the original data, three constructed variables are included in the data set: the three-category classification variable utilized to construct Figure 1, our index of cosmopolitantism-traditionalism used to rank school districts from least to most conservative, and a code indicating the state in which the school is located.

Consistent with practices to protect the confidentiality of responses, and conforming to IRB requirements, we do not include geographic codes that identify the specific school or school district. Likewise, the index of cosmopolitantism-traditionalism is rounded to a single decimal so that it is impossible to reverse-engineer the original geographic data. Researchers interested in finer geography should contact the corresponding author; an IRB approved data security plan will be required for release of detailed geographic identifiers.

## References


S1. D. A. Dillman, *Mail and Internet Surveys: The Tailored Design Method* (Wiley, New York, 2000).

S2. A. H. Mokdad, D. F. Stroup, W. H. Giles; Behavioral Risk Factor Surveillance Team, Public health surveillance for behavioral risk factors in a changing environment. Recommendations from the Behavioral Risk Factor Surveillance Team. *MMWR Recomm. Rep.* **52**, 1 (2003).

S3. American Association for Public Opinion Research (AAPOR), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (AAPOR, Deerfield, IL, ed. 6, 2009).

S4. L. Biener, C. A. Garrett, E. A. Gilpin, A. M. Roman, D. B. Currivan, Consequences of declining survey response rates for smoking prevalence estimates. *Am. J. Prev. Med.* **27**, 254 (2004).

S5. B. C. Burden, Voter turnout and the National Election Studies. *Polit. Anal.* **18**, 389 (2000).

S6. E. Singer, Introduction: Nonresponse bias in household surveys. *Public Opin. Q.* **70**, 637 (2006).

S7. S. M. Elan, Differences between educators and the public on questions of education policy. *Phi Delta Kappan* **69**, 294 (1987).

S8. C. A. Langdon, N. Vesper, The sixth Phi Delta Kappa poll of teachers' attitudes toward the public schools. *Phi Delta Kappan* **81**, 607 (2000).

S9. F. M. Hess, R. Maranto, S. Milliman, Resistance in the trenches: What shapes teachers' attitudes toward school choice? *Educ. Policy* **14**, 195 (2000).

S10. R. M. Groves *et al*., *Survey Methodology* (Wiley, New York, 2004).

S11. S. Keeter, C. Miller, A. Kohut, R. Groves, S. Presser, Consequences of reducing nonresponse in a national telephone survey. *Public Opin. Q.* **64**, 125 (2000).

S12. D. Merkle, M. Edelman, in *Survey Nonresponse*, Robert Groves, Don Dillman, John Eltinge, Rod Little, Eds. (Wiley, New York, 2002), pp. 243–58.

S13. L. Kish, *Survey Sampling* (Wiley, New York, 1965).

S14. S. Lohr, *Sampling: Design and Analysis* (Duxbury, New York, 1999).

S15. M. B. Berkman, E. Plutzer, *Evolution, Creationism, and the Battle to Control America's Classrooms* (Cambridge Univ. Press, Cambridge, 2010).