Public Opinion Quarterly, Vol. 74, No. 2, Summer 2010, pp. 328–343

MEASURING VOTER TURNOUT BY USING THE RANDOMIZED RESPONSE TECHNIQUE
EVIDENCE CALLING INTO QUESTION THE METHOD'S VALIDITY

ALLYSON L. HOLBROOK*
JON A. KROSNICK*

Abstract Surveys usually yield reported rates of voting in elections that are higher than official turnout figures, a phenomenon often attributed to intentional misrepresentation by respondents who did not vote and would be embarrassed to admit that. The experiments reported here tested a procedure for reducing social desirability response bias by allowing respondents to report secretly whether they voted: the "randomized response technique." In a national telephone survey of a sample of American adults and eight national surveys of American adults conducted via the Internet, respondents were either unable or unwilling to implement the randomized response technique properly, raising questions about whether this technique has ever worked properly to achieve its goals.

Introduction

Self-reports in surveys often overestimate voter turnout (e.g., Bernstein, Chadha, and Montjoy 2001; McDonald 2003), and researchers have argued that this occurs partly because some respondents intentionally misreport that

ALLYSON L. HOLBROOK is Associate Professor of Public Administration and Psychology at the Survey Research Laboratory of the University of Illinois at Chicago, Chicago, IL, USA. JON A. KROSNICK is Professor of Communication, Political Science, and Psychology at Stanford University, Stanford, CA, USA, and University Fellow at Resources for the Future. The data for study 3 were collected through Time-Sharing Experiments for the Social Sciences (TESS), supported by a grant [SES-0094964 to Diana Mutz and Arthur Lupia] from the National Science Foundation. The data for studies 1, 2, and 4 were collected by the Stanford Institute for the Quantitative Study of Society. The authors thank Norman Nie, Douglas Rivers, and LinChiat Chang for their roles in making studies 1, 2, and 4 possible. *Address correspondence to Allyson L. Holbrook, 412 S. Peoria Street (MC336), Chicago, IL 60607, USA; e-mail: allyson@uic.edu or Jon A. Krosnick, Rm.434 McClatchy Hall, 450 Serra Mall, Stanford, California, 94305, USA; e-mail: krosnick@stanford.edu.

doi: 10.1093/poq/nfq012 Advance Access publication March 26, 2010

they voted to present themselves favorably (Corbett 1991; Lyons and Scheb 1999; Aarts 2002; Brockington and Karp, 2002; Andolina et al. 2003; Lutz 2003; Blais et al. 2004; see Holbrook and Krosnick 2010 for a review). Although some previous attempts to reduce over-reporting by reducing social desirability pressures have been unsuccessful (e.g., Presser 1990; Abelson, Loftus, and Greenwald 1992), Holbrook and Krosnick (2010) recently showed that the Item Count Technique (ICT) reduced over-reporting in a telephone survey (though not in Internet surveys), suggesting that social desirability response bias may distort telephone reports of turnout. In this article, we explore whether another technique, the randomized response technique (RRT), can be used to reduce social desirability response bias in turnout reports.

THE RANDOMIZED RESPONSE TECHNIQUE

The RRT allows respondents to report an embarrassing fact in a completely secret way, so they have no motivation to lie (e.g., Warner 1965; Himmelfarb and Lickteig 1982; Lensvelt-Mulders et al. 2005). If respondents who implement the RRT acknowledge possessing a socially undesirable attribute more often than respondents who report that attribute directly, the magnitude of this difference reveals the proportion of the latter group who intentionally lied and the size of the social desirability response bias.

The RRT can be implemented in many different ways (see Fox and Tracy 1986). Most RRT research has been done in face-to-face surveys (e.g., Locander, Sudman, and Bradburn 1976), but it has also been used in telephone surveys (e.g., Orwin and Boruch 1982; Stem and Steinhorst 1984), Internet surveys (e.g., Musch, Bröder, and Klauer 2001; Lensvelt-Mulders et al. 2006), and mail surveys (e.g., Stem and Steinhorst 1984).

Many studies have suggested that the RRT worked effectively (see Fox and Tracy 1986; Lensvelt-Mulders et al. 2005). In contrast to respondents answering direct questions, respondents answering RRT questions were significantly more likely to report a variety of socially undesirable behaviors: cheating (Shotland and Yankovski 1982; Scheers and Dayton 1987; Franklin 1989; Kerkvliet 1994), falsifying income tax reports (Himmelfarb and Lickteig 1982; Musch, Bröder, and Klauer 2001), stealing (Franklin 1989; Wimbush and Dalton 1997), behaving unprofessionally (Buchman and Tracy 1982), using illegal drugs (Goodstadt and Gruson 1975; Goodstadt and Cook 1978; Himmelfarb and Lickteig 1982), drinking alcohol and smoking cigarettes by high school students (Barth and Sandler 1976; Fisher, Kupferman, and Lesser 1992), smuggling liquor in Norway (Nordlund, Holme, and Tamsfoss 1994), alcohol abuse (Volicer and Volicer 1982; Volicer et al. 1983), having had an abortion or being willing to consider having one (I-Cheng, Chow, and Rider 1972; Krotke and Fox 1974; Rider et al. 1976; Tezcan and Omran 1981; Himmelfarb and Lickteig 1982; Lara et al. 2004), enjoying pornography

(Himmelfarb and Lickteig 1982), performing possibly stigmatized sexual behaviors (Fidler and Kleinknecht 1977; Franklin 1989), failing to report a death to the government (Madigan et al. 1976), being at least slightly likely to perform date rape if there was no possibility of being caught (Himmelfarb 2004), being prejudiced (Jarman 1992; Himmelfarb 2004), and being depressed (Himmelfarb 2004). The less admirable the attribute, the more the RRT has increased prevalence estimates relative to answers to direct questions (Fidler and Kleinknecht 1977; Himmelfarb and Lickteig 1982).[1] Several studies reported evidence suggesting that RRT estimates were more accurate: They more closely matched rates from official records than did rates from direct self-reports (Horvitz, Shah, and Simmons 1967; Liu and Chow 1976; Lamb and Stem 1978; van der Heijden et al. 2000; see also Lensvelt-Mulders et al. 2005).

However, other studies have found no significant differences between RRT and direct self-report measures (Locander, Sudman, and Bradburn 1976; Bégin and Boivin 1980; Tamhane 1981; Akers et al. 1983; Danermark and Swensson 1987; Duffy and Waterton 1988; Franklin 1989). And direct self-reports sometimes yielded significantly higher estimates of socially undesirable behaviors than did the RRT (Wiseman, Moriarty, and Schafer 1975; Brewer 1981; Beldt, Daniel, and Garcha 1982; Williams and Suen 1994).

Furthermore, some studies yielded evidence suggesting that some respondents did not follow RRT procedures properly (Edgell, Himmelfarb, and Duchan 1982; Fox and Tracy 1986; Mangat and Singh 1990**;** Clark and Desharnais 1998; Böckenholt and van der Heijden 2007). For example, when answering an RRT question designed to assess having had homosexual experiences, about 25 percent of respondents did not follow instructions and answered "no" when they were randomly assigned to have answered "yes" (Edgell, Himmelfarb, and Duchan 1982). Similarly, by assigning groups of respondents to RRT procedures with different randomization probabilities, Clark and Desharnais (1998) found that some respondents did not answer as instructed. And Stem and Steinhorst (1984) reported that the RRT procedure had to be described several times to many of their telephone survey respondents before they could implement it properly. Thus, there are some reasons to hesitate about the efficacy of the RRT.

VOTER TURNOUT REPORTS

Only one study has used the RRT to study voter turnout. Members of a small sample of Chicago residents interviewed face-to-face were randomly assigned to report turnout in a primary election either directly (N=80) or via the RRT

1. Other studies found evidence supportive of the RRT, but did not report tests of statistical significance (Shimizu and Bonham 1978; Zdep et al. 1979; Scheers and Dayton 1982; Stem and Steinhorst 1984; Soeken and Macready 1986; Weissman, Steer, and Lipton 1986).

($N = 50$; Locander, Sudman, and Bradburn 1976). Over-reporting relative to official records was equivalent across the two groups, suggesting that social desirability bias did not distort reports in this mode.

We tested whether the RRT reduces estimates of turnout in a telephone survey of a representative national sample of American adults (study 1), in Internet surveys of representative national samples (studies 2 and 3), and in six Internet surveys conducted with non-probability samples (study 4).

Methods[2]

STUDY 1

Respondents: A representative national random digit dial (RDD) sample of 966 American adults was interviewed by telephone by Schulman, Ronca, and Bucuvalas, Inc. (AAPOR RR3 = 35.6 percent).

Conditions: Twenty percent of respondents were randomly assigned to be asked the traditional American National Election Studies (ANES) turnout question ($N = 176$): "In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. How about you—did you vote in the presidential election held on November 7, 2000?"[3]

Twenty percent of respondents were randomly assigned to RRT condition 1 ($N = 189$), and another 20 percent were randomly assigned to RRT condition 2 ($N = 179$; see Folsom et al. 1973).[4] Respondents in RRT condition 1 were asked, "The next question deals with whether or not you voted in the presidential election held on November 7, 2000. People are sometimes reluctant to say whether or not they voted because it's a personal question. In order to respect your privacy, this question uses a method to keep your answer totally confidential, so even I cannot tell for sure whether you voted or not. To begin, would you please get a coin to use to answer the next question?" Interviewers then paused and asked, "Do you have a coin?" Respondents who answered "no" were asked, "Would you please get a coin to use to answer the next

2. More methodological details about each data collection are provided in the Appendix. In all four studies, some respondents were assigned to experimental conditions not described in this article, so the sample sizes for the conditions described in this article are less than the total survey samples.

3. The ANES turnout question wording varied over the years; this wording was employed in 1952–1960, 1964–1998, and 2002.

4. One respondent did not provide a substantive response to the direct question and was excluded from the analyses. Thirty-eight and twenty respondents did not provide substantive answers in RRT conditions 1 and 2, respectively, and were excluded. An additional 422 respondents were assigned to experimental conditions not described in this article.

question? You can't answer it without having a coin," and the interviewer waited until the respondent said he or she had a coin.

Respondents were then told, "First, please flip the coin. Look at the side facing up, but DON'T tell me which side is facing up. If the coin landed *tails* up, when I say 'now please answer' in a moment, please answer this question: 'Did you watch a news program on television at 11:00 last night?' If the coin landed *heads* up, when I say 'now please answer' in a moment, please answer this question: 'Did you vote in the presidential election held on November 7, 2000?' Now please answer." These respondents were then also asked, "Now, a question on a different topic: Do you have any older living siblings?"

Respondents in RRT condition 2 were read the same introduction and were then told, "If the coin landed *tails* up, when I say 'now please answer' in a moment, please answer this question: 'Do you have any older living siblings?' If the coin landed *heads* up, when I say 'now please answer' in a moment, please answer this question: 'Did you vote in the presidential election held on November 7, 2000?' Now please answer." These respondents were then also asked, "Now, a question on a different topic: Did you watch a news program on television at 11:00 last night?"

The proportion of respondents who voted in each condition was calculated as $[y - ((N/2)*z)] * (2/N)$, where y is the number of "yes" responses, z is the proportion of respondents who said they performed the nonsensitive behavior when asked directly, and N is the number of respondents in the condition. The standard error of this estimate is the square root of $(1/P^2) * [P_Y*(1-P_Y)/N_c + (1-P)^2*P_{NS2}*(1-P_{NS2})/N_o]$, where P is the probability of being asked the nonsensitive question, P_Y is the proportion of "yes" responses in the condition, N_c is the condition sample size, P_{NS2} is the proportion of people who said "yes" to the direct self-reported nonsensitive question in the other condition, and N_o is the sample size in the other condition.

STUDY 2

Respondents: Study 2's survey was administered by Knowledge Networks to a representative national sample of American adults recruited via RDD telephone interviews (see Knowledge Networks 2006). Of the 1,533 adults sent an email invitation to participate in our study, 1,175 (77 percent) did so. The cumulative response rate (taking into account the household recruitment rate into the panel, completion of profiles by members of recruited households, and break-offs among respondents asked to complete the survey) was 15.3 percent.

Conditions: Ten percent of the respondents were randomly assigned to be asked the traditional ANES turnout question ($N = 117$). Twenty percent were

randomly assigned to RRT condition 1 (N = 234), and 20 percent were assigned to RRT condition 2 (N=238). Procedures and instructions were adapted from study 1 for Internet administration (see Appendix).[5]

STUDY 3

Respondents: Study 3's survey was administered by Knowledge Networks; 9,896 adults were invited to participate, and 6,094 (62 percent) did so. The cumulative response rate (calculated as in study 2) was 18.2 percent.

Conditions: One-third of respondents were randomly assigned to be asked the traditional ANES direct self-report turnout question (N = 2,018).

Another one-third of respondents (N = 2,012) were asked to get a coin and flip it. People whose coins came up tails were told to select the radio button marked "no." People whose coins came up heads were asked to select the appropriate radio button (marked "yes" or "no") to answer the question "Did you vote in the elections held on November 5, 2002?"[6]

The proportion of people who voted was calculated as ("Yes" responses)/((Total N)/2). The standard error of this estimate is the square root of $P_Y(1-P_Y)/(N*P^2)$, where P_Y is the proportion of "yes" responses, N is the sample size, and P is the likelihood the respondent will be asked the sensitive question (Horvitz, Shah, and Simmons 1967).

STUDY 4

Study 4 implemented the RRT in six Internet surveys conducted by six companies that collect data from panels of people who volunteered to do surveys: Gozing (N = 1,075), Greenfield Online (N = 1,223), Harris Interactive (N = 2,406), SPSS (N = 1,112), SSI (N = 1,103), and Survey Direct (N = 1,086). Each company was asked to collect data from a group of people as similar to a representative national sample as possible (see the Appendix for details).

Results

The proportions of respondents who were asked the traditional ANES direct self-report question and who reported voting in the 2000 election were 72.0 percent (study 1), 66.1 percent (study 2), and 69.9 percent (study 4); 59.5

5. Two respondents did not provide substantive responses to the direct question and were excluded. Sixteen and nineteen respondents did not provide substantive responses in RRT conditions 1 and 2, respectively, and were excluded. An additional 584 respondents were assigned to experimental conditions not described in this article.

6. Thirty-five respondents declined to answer a turnout question (we do not know which question version each of these people were asked) and were excluded. An additional 2,029 respondents were assigned to experimental conditions not described in this article.

Table 1. Estimates of Turnout

	Study 1	Study 2	Study 3	Study 4
Direct self-reports				
Traditional NES wording				
Number "YES"	126	76	1,200	599
Sample Size	175	115	2,018	857
Reported turnout	**72.0%**	**66.1%**	**59.5%**	**69.9%**
RRT				
One-condition design				
Number "YES" (y)			1,046	
Sample size			2,012	
Estimated turnout			**102.0%**	
Standard deviation of estimate			**2.2%**	
95% confidence interval			**97.6–106.4%**	
Two-condition design				
Condition 1				
Number "YES" (*y*)	110	115		953
News program (*z*)	.24	.20		.37
Sample size (*N*)	151	217		1,493
Estimated turnout	**121.7%**	**86.0%**		**90.4%**
Standard deviation estimate	**7.7%**	**7.3%**		**85.0–95.9%**
95% confidence interval	**106.1–137.3%**	**71.7–100.3%**		
Condition 2				
Number "YES" (*y*)	121	153		1,031
Older living sibling (*z*)	.50	.63		.52
Sample size (*N*)	159	219		1,521
Estimated turnout	**102.2%**	**77.7%**		**83.6%**
Standard deviation estimate	**7.9%**	**7.0%**		**2.6%**
95% confidence interval	**86.7–117.7%**	**64.0–91.5%**		**78.3–88.9%**
Conditions 1 + 2				
Number "YES" (*y*)	231	268		1,984
Sample size	310	436		3,014
Estimated turnout[a]	**111.6%**	**81.3%**		**87.1%**

[a]These are averages of the estimated turnout in conditions A and B, weighted by sample size.
NOTE.—Sample sizes reported in this table indicate the number of valid responses.

percent of them reported voting in the 2002 election (study 3; see table 1). The RRT consistently yielded turnout rates *higher* than did the direct self-reports, and often greater than 100 percent, which is impossible and suggests that respondents did not accurately implement the technique (study 1: 111.6 percent across conditions; study 2: 81.3 percent across conditions; Study 3: 102.0 percent; Study 4: 87.1 percent across conditions).[7]

It is theoretically possible that respondents properly implemented the RRT and that their coin flips came up "heads" frequently enough to produce the above percentages. To do so, 112 percent, 147 percent, 86 percent, and 93 percent of respondents' coins would have had to come up "heads" in studies 1–4, respectively. These figures for studies 1 and 2 are impossible. In studies 3 and 4, the chances that these percentages of "heads" would have been observed are incredibly small (comparison to rate of .50: study 3: $\chi^2(1) = 599$, $p < .001$; study 4: $\chi^2(1) = 1370$, $p < .001$).

Consistent with prior research suggesting implementation problems with the RRT, more respondents elected not to answer those questions than answered the direct self-report questions. In studies 1 and 2, small percentages of respondents failed to answer the direct self-report questions (.6 percent and 1.7 percent, respectively), whereas notably larger proportions of people failed to answer the RRT questions (15.8 percent and 7.6 percent, respectively).[8] In theory, these larger non-response rates could account for some or all of the observed differences in reported turnout rates between the direct self-report and RRT questions if all the people who declined to answer the RRT questions did not, in fact, vote and would have said so in response to an RRT question. But even if that were true, the RRT estimates of turnout would still have been larger than those yielded by the direct questions: 88.9 percent in study 1 and 76.7 percent in study 2.

One might imagine that the implausible results from the RRT occurred because respondents with more limited cognitive skills were more likely to make mistakes when attempting to carry out the instructions. But in fact, in study 3, for example, implausible differences in turnout rates between the direct self-report and RRT questions were largest among the *most* educated respondents (people with a college degree or more: $\Delta = 51$ percentage points; RRT = 130 percent; direct = 79 percent) and smallest among the least educated respondents (people with less than a high school degree: Δ=28 percentage

7. Of the people asked the nonsensitive questions directly in studies 1, 2, and 4, the proportions answering affirmatively ranged from 20 percent to 63 percent (see table 1). Although these proportions impact the prevalence of voting estimated in the RRT, the prevalence of the nonsensitive behavior was unassociated with the extent to which the RRT estimated higher turnout than the direct self-report condition.

8. We were not able to make these comparisons using the data from studies 3 and 4, because we were not told which question version had been asked of each respondent who failed to answer the turnout questions.

points; RRT = 72 percent; direct = 44 percent). This suggests that the failure of the RRT was not due to the cognitive difficulty of the task.[9]

Discussion

The RRT was a failure in both the telephone and Internet surveys, yielding results in the opposite direction to what was expected. One possible explanation for these findings and those of most past RRT studies is that the RRT increases reports of all attributes, regardless of their social desirability, and does not increase measurement accuracy. Past researchers have never confirmed whether individual reports of attributes are more accurate when reported via the RRT (because such checking is impossible) and have simply presumed that more reports of undesirable attributes indicate more accuracy. We found that the RRT increased reports of a socially desirable attribute as well, which cannot be due to the removal of social desirability pressures and resulting increased accuracy. Consistent with this reasoning, among the few studies that have compared RRT and direct self-report estimates of socially admirable attributes, none yielded consistent evidence that the RRT significantly reduced reported rates (e.g., Locander, Sudman, and Bradburn 1976; Bégin, Boivin, and Bellerose 1979; Himmelfarb and Lickteig 1982; Williams and Suen 1994; Berrens, Bohara, and Kerkvliet 1997). This calls into question interpretations of all past RRT studies and raises serious questions about whether the RRT has practical value for increasing survey reporting accuracy.

Our studies involved data collected via the telephone and the Internet. Past research suggests that social desirability pressures are minimal during Internet questionnaire completion (e.g., Holbrook and Krosnick 2010), so it may come as no surprise that the RRT did not reduce turnout reports in this mode. But the finding that the RRT increased turnout reports to impossible levels means that this measurement approach is not only unnecessary in this mode but is in fact deleterious. In contrast, social desirability pressures do inflate turnout reports during telephone interviews (Holbrook, Green, and Krosnick 2003; Holbrook and Krosnick 2010). Yet the RRT did not reduce turnout reports in that mode either, and again inflated calculated turnout rates to implausible levels. This failure is not attributable to lack of social desirability pressures.

Most past published RRT studies involved face-to-face data interviewing (see Orwin and Boruch 1982; Stem and Steinhorst 1984 for reviews), so perhaps the RRT works better in face-to-face interviews than in telephone and Internet data collection. This mode may allow the interviewer to supervise the implementation of the randomization procedure (e.g., a coin flip) more than occurs in other modes. So perhaps it is implemented properly more often,

9. Reliable RRT estimates could not be calculated within education subgroups in studies 1 and 2 because of smaller sample sizes. Study 4's nonprobability samples included almost no low education respondents, so we could not calculate reliable RRT estimates for these individuals.

yielding more accurate reports. But even in face-to-face interviews, the interviewer cannot supervise the cognitive processes that must be implemented properly in respondents' heads. And even in this mode, past studies have consistently found that the RRT did not reduce reports of socially desirable behavior (e.g., Locander, Sudman, and Bradburn 1976; Himmelfarb and Lickteig 1982). Therefore, we suspect that the failure of the RRT seen in the present studies is not attributable to the modes used. We therefore look forward to future studies testing whether the RRT does, in fact, lead to more accurate reporting in any mode or whether it simply increases rates of all reported attributes due to inevitable implementation errors.

Appendix: Methodological Details

STUDY 1

Procedures: Interviewing was done between June 15, 2004, and September 16, 2004. Of 6,990 initial phone numbers in the sample, reverse lookup procedures identified addresses for 2,518 of these numbers and pre-notification letters were sent to these addresses (36 percent of the sample). Up to 12 attempts were made to contact each number, and one refusal conversion attempt was made for each number if needed. A non-contact letter was mailed to 879 households for which addresses could be obtained but who had not completed the survey partway through the field period. Letters were also sent to 95 households for whom addresses were available in order to convert refusals to completed interviews. The non-contact and refusal conversion letters offered an incentive of $10 for completing the survey.

STUDY 2

Procedures: Knowledge Networks recruited panel members through random digit dialing (RDD) telephone interviewing. This sampling technique entailed coverage error due to excluding households without working landline telephones (about five percent of the country's population) and individuals who were homeless or lived in institutions, but this approach provided significantly broader coverage than sampling techniques that draw only from regular computer users with Internet access and has the same limitations as telephone surveys in this regard.

Before the initial telephone calls were made, households for which Knowledge Networks was able to recover a valid postal address (about 70 percent of the RDD sample) were sent letters describing the proposed exchange relationship. Specifically, they were told that they had been randomly selected to participate in the survey panel, they would not pay any cost incurred, confidentiality was assured, and a Knowledge Networks staff member would call them within a week of receipt of the letter. During the initial RDD telephone

interview, respondents were told they had been selected to participate in an important national study. Households without Internet access were told that KN would give them an Internet appliance and an Internet service connection in exchange for their participation in surveys. Potential panel members who had access to the Internet were asked to use their own equipment and were given points for participation that could be redeemed for cash. Once a household had enrolled in the panel, respondents were asked to answer profile surveys that measured many attributes.

Emails inviting respondents to complete our survey were sent on June 18, 2004, and no responses were accepted after July 2, 2004.

STUDY 3

Procedures: Knowledge Networks sent emails inviting panel members to complete our survey on November 15, November 20, and November 26, 2002 (emails were sent to different groups of respondents on these three days), and no responses were accepted after December 5, 2002.

STUDY 4

Survey data for study 4 were collected via the Internet by six different companies using panels of volunteer respondents. Each organization was asked to provide a survey sample that was representative of adults from the 50 U.S. states. The organizations used somewhat different methodologies, described below. For organizations that maintained a panel of respondents, we computed the cooperation rate by dividing the number of panel members who fully or partially completed the questionnaire by the number of panel members who were invited to do so, which is equivalent to AAPOR's Cooperation Rate 2.

Firm 1: Firm 1 collected data from members of a panel of approximately 2.2 million people. Panel members opted into the panel via the Internet and were recruited via many methods through Firm 1's affiliates, including text links in newsletters, banner ads, email invitations, and word of mouth. On average, respondents had been in the panel for six months. They were invited to complete no more than one survey every week. Panel attrition was approximately 30 percent per year. Respondents were given $1 for completing a questionnaire that could be obtained in cash via Paypal or buy.com or could be used to pay for music downloads. Panel members were invited to participate in our survey in proportions matching quotas (reflecting Census estimates of the population) for household income, ethnicity, education, and gender. Emails inviting 2,123 people to complete our questionnaire were sent on July 26, 2004, and no responses were accepted after August 1, 2004; 1075 (51 percent) people completed the questionnaire.

Firm 2: Firm 2 collected data via the Internet from members of a panel of over 1.7 million respondents. Panel members opted into the panel via the Internet and were recruited via high-value media including embedded text links, editorial inclusion, targeted opt-in email lists, word of mouth, co-registrations, and online promotions. On average, members had been in the panel for 18 months. They completed no more than one survey every two weeks and were invited to participate in one or two surveys per week. Panel attrition was approximately 32 percent per year. Respondents were given a chance to enter a weekly drawing for a $3,000 prize as compensation for questionnaire completion. Panel members were invited to complete our questionnaire in proportions matching quotas for gender, age, and region. Emails inviting people to complete the survey were sent on June 11, 2004, and no responses were accepted after June 14, 2004. Firm 2 invited a total of 50,000 panel members to participate, and 1,223 (2.4 percent) did so.

Firm 3: Firm 3 conducts surveys via the Internet with members of a panel of approximately 5 million panel members in the United States, who were recruited for the panel through Internet signups. Panel members were typically invited to participate in two to three surveys per month. Respondents were given 100 points to redeem for prizes and they were entered in a monthly $10,000 sweepstakes as compensation for completing questionnaires. Panel members were invited to participate in our survey in proportions matching quotas for age, gender, and region, with additional efforts to select Hispanics and African Americans to mirror the U.S. population. Emails inviting respondents to complete the questionnaire were sent on June 11, 2004, and no responses were accepted after June 21, 2004. Email invitations were sent to 11,530 veteran panel members and 45,014 new panel members (who had not completed a prior survey); 1841 of the former group (16 percent) and 565 of the latter (1.5 percent) did so.

Firm 4: Firm 4 recruited respondents to do Internet surveys through advertisements that appeared on ISP Web sites inviting visitors to participate in a survey. Respondents who clicked on an advertisement asking for survey participants were first asked a series of screening demographic questions. Based on this information, potential respondents were assigned to participate in one of a series of surveys via quota sampling in proportions to match the population on age, gender, income, and region. A respondent could participate in only two surveys per month and only one survey on a particular topic every 90 days. Respondents were offered either a $4.50 credit on their monthly ISP bill or 300 airline frequent flier miles for completing a survey. A total of 1,112 respondents completed the survey between June 16, 2004, and July 1, 2004. Given the recruitment methodology, a cooperation rate cannot be calculated, because the number of people who were "invited" to complete the survey via the banner ads is not known.

Firm 5: Internet survey data were collected from respondents in a sample provided by Firm 5, whose panel included approximately 1.6 million members. Panel members were recruited via the Internet, RDD invitations, referrals, and banner ads. Panel members completed one or two surveys per month, and respondents received no more than one or two emails per week inviting them to respond. Panel attrition was approximately 20 to 25 percent per year, and panelists had participated, on average, for one year. Panel members were invited to complete our questionnaire in proportions matching 2001 CPS estimates for gender, age, and income. People were offered a chance to enter the monthly prize pool to win one of 114 prizes worth $10,000 in exchange for completing our questionnaire. Emails inviting respondents to complete the questionnaire were sent on June 23, 2004, and no responses were accepted after June 30, 2004. Of the 9,921 panel members invited to participate, 1,103 (11.1 percent) did so.

Firm 6: Firm 6 maintained an Internet survey panel of 2.5 million potential respondents, who were recruited through more than 400 Web sites. Approximately 50,000 people joined their panel per month. Panel members were typically sent no more than four to six invitations to participate in a survey per month. On average, people were on the panel for a total of 18 months. No incentive was offered for completing our survey. The sample of panel members selected was drawn to match the population on age, gender, and geography, but no formal quotas were used. Emails inviting respondents to complete the questionnaire were sent on August 25, 2004, and no responses were accepted after September 1, 2004. Firm 6 invited a total of 14,000 panel members to participate, and 1086 (8 percent) did so.

References

Aarts, Kees. 2002. *Electoral Turnout in West-European Democracies.* Paper presented at the Annual Meeting of the American Political Science Association, Boston, MA, USA.

Abelson, Robert P., Elizabeth F. Loftus, and Anthony G. Greenwald. 1992. "Attempts to Improve the Accuracy of Self-Reports of Voting." In *Questions about Questions*, ed. Judith M. Tanur. 138–53. New York: Russell Sage Foundation.

Akers, Ronald L., James Massey, William Clarke, and Ronald M. Lauer. 1983. "Are Self-Reports of Adolescent Deviance Valid? Biochemical Measures, Randomized Response, and the Bogus Pipeline in Smoking Behavior." *Social Forces* 62:234–51.

Andolina, Molly, Scott Keeter, Cliff Zukin, and Krista Jenkins. 2003. *A Guide to the Index of Civic and Political Engagement*. College Park, MD: Center for Information and Research on Civic Learning and Engagement.

Barth, Jeremy T., and Howard M. Sandler. 1976. "Evaluation of the Randomized Response Technique in a Drinking Survey." *Journal of Studies on Alcohol* 37:690–3.

Bégin, Guy, and Michel Boivin. 1980. "Comparison of Data Gathered on Sensitive Questions via Direct Questioning, Randomized Response Technique, and a Projective Method." *Psychological Reports* 47:743–50.

Bégin, Guy, Michel Boivin, and J. Bellerose. 1979. "Sensitive Data Collection through the Randomized Response Technique: Some Improvements." *Journal of Psychology* 101:53–65.

Beldt, Sandra F., Wayne W. Daniel, and Bikramjit S. Garcha. 1982. "The Takahasi-Sakasegawa Randomized Response Technique." *Sociological Methods and Research* 11:101–11.

Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. "Over-Reporting Voting: Why It Happens and Why It Matters." *Public Opinion Quarterly* 65:22–44.

Berrens, Robert P., Alok K. Bohara, and Joe Kerkvliet. 1997. "A Randomized Response Approach to Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics Association* 79:252–66.

Blais, André, Elisabeth Gidengil, Neil Nevitte, and Richard Nadeau. 2004. "Where Does Turnout Decline Come From?" *European Journal of Political Research* 43(2):221–36.

Böckenholt, Ulf, and Peter G. M. Van der Heijden. 2007. "Item Randomized-Response Models for Measuring Noncompliance: Risk-Return Perception, Social Influences, and Self-Protective Responses." *Psychometrika* 72:245–62.

Brewer, K. R. W. 1981. "Estimating Marihuana Usage Using Randomized Response-Some Paradoxical Findings?" *Australian Journal of Statistics* 23:139–48.

Brockington, David, and Jeffrey Karp. 2002. *Social Desirability and Response Validity: A Comparative Analysis of Over-Reporting Turnout in Five Countries.* Paper presented at the Annual Meeting of the American Political Science Association, Boston, MA, USA.

Buchman, Thomas A., and John A. Tracy. 1982. "Obtaining Responses to Sensitive Questions: Conventional Questionnaire versus Randomized Response Technique." *Journal of Accounting Research* 20:263–71.

Clark, Stephen J., and Robert A. Desharnais. 1998. "Honest Answers to Embarrassing Questions: Detecting Cheating in the Randomized Response Model." *Psychological Methods* 3:160–8.

Corbett, Michael. 1991. *American Public Opinion*. White Plains, NY: Longman.

Danermark, Berth, and Bengt Swensson. 1987. "Measuring Drug Use among Swedish Adolescents." *Journal of Official Statistics* 3:439–48.

Duffy, John C., and Jennifer J. Waterton. 1988. "Randomised Response vs. Direct Questioning: Estimating the Prevalence of Alcohol-Related Problems in a Field Survey." *Australian Journal of Statistics* 30(1):1–14.

Edgell, Stephen E., Samuel Himmelfarb, and Karen L. Duchan. 1982. "Validity of Forced Response in a Randomized Response Model." *Sociological Methods and Research* 11: 89–110.

Fidler, Dorothy S., and Richard E. Kleinknecht. 1977. "Randomized Response Versus Direct Questioning: Two Data-Collection Methods for Sensitive Behavior." *Psychological Bulletin* 84:1045–9.

Fisher, Martin, Linda B. Kupferman, and Martin Lesser. 1992. "Substance Use in a School-Based Clinic Population: Use of the Randomized Response Technique to Estimate Prevalence." *Journal of Adolescent Health* 13:281–5.

Folsom, Ralph E., Bernard G. Greenberg, Daniel G. Horvitz, and James R. Abernathy. 1973. "The Two Alternative Questions Randomized Response Model for Human Surveys." *Journal of the American Statistical Association* 68:525–30.

Fox, James A., and Paul E. Tracy. 1986. *Randomized Response: A Method for Sensitive Surveys*. Beverly Hills, CA: Sage.

Franklin, LeRoy A. 1989. "Randomized Response Sampling from Dichotomous Populations with Continuous Randomization." *Survey Methodology* 15:225–35.

Goodstadt, Michael S., and Gaynoll Cook. 1978. "The Validity of Reported Drug Use: The Randomized Response Technique." *International Journal of the Addictions* 13:359–67.

Goodstadt, Michael S., and Valerie Gruson. 1975. "The Randomized Response Technique: A Test on Drug Use." *Journal of the American Statistical Association* 70:814–8.

Himmelfarb, Samuel. 2004. The Multi-Item Randomized Response Technique. Unpublished manuscript, University of Louisville, KY.

Himmelfarb, Samuel, and Carl Lickteig. 1982. "Social Desirability and the Randomized Response Technique." *Journal of Personality and Social Psychology* 43:710–7.

Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. "Telephone vs. Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias." *Public Opinion Quarterly* 67:79–125.

Holbrook, Allyson L., and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74:37–67.

Horvitz, Daniel G., B. V. Shah, and Walt R. Simmons. 1967. "The Unrelated Randomized Response Model." Proceedings of the American Statistical Association, Social Statistics Section, 65–72.

I-Cheng, Chi, L. P. Chow, and Rowland V. Rider. 1972. "The Randomized Response Technique as Used in the Taiwan Outcome of Pregnancy Study." *Studies in Family Planning* 3:265–9.

Jarman, Bonita J. 1992. *Racism and Social Desirability: Exploring the Multi-Item Randomized Response Technique*. M.A. thesis, University of Louisville, KY.

Kerkvliet, Jan. 1994. "Cheating by Economics Students: A Comparison of Survey Results." *Journal of Economic Education* 25:121–33.

Knowledge Networks. 2006. *Knowledge Networks Remains Only Online Research Firm to Deliver Six Elements of Accuracy.* Available at http://www.knowledgenetworks.com/news/releases/2006/100406_knpanel.htm.

Krotke, K. J., and B. Fox. 1974. "The Randomized Response Technique, the Interview, and the Self-Administered Questionnaire: An Empirical Comparison of Fertility Reports." In Proceedings of the American Statistical Association, Social Statistics Section, 367–71.

Lamb, Charles W., and Donald E. Stem. 1978. "An Empirical Validation of the Randomized Response Technique." *Journal of Marketing Research* 15:616–21.

Lara, Diana, Jennifer Stricker, Claudia Diaz Olaverrieta, and Charlotte Ellertson. 2004. "Measuring Induced Abortion in Mexico." *Sociological Methods and Research* 32:529–58.

Lensvelt-Mulders, Gerty J. L. M., Joop J. Hox, Peter G. M. van der Heijden, and Cora J. M. Maas. 2005. "Meta-Analysis of Randomized Response Research." *Sociological Methods and Research* 33(3):319–48.

Lensvelt-Mulders, Gerty J. L. M., Peter G. M. van der Heijden, Olav Laudy, and Ger van Gils. 2006. "A Validation of a Computer-Assisted Randomized Response Survey to Estimate the Prevalence of Fraud in Social Security." *Journal of the Royal Statistical Society Series A* 169:305–18.

Liu, P. T., and L. P. Chow. 1976. "The Efficiency of the Multiple Trial Randomized Response Technique." *Biometrics* 32:607–18.

Locander, William, Seymour Sudman, and Norman Bradburn. 1976. "An Investigation of Interview Method, Threat and Response Distortion." *Journal of the American Statistical Association* 71:269–75.

Lutz, George. 2003. *Participation, Cognitive Involvement, and Democracy: When Do Low Turnout and Low Cognitive Involvement Make a Difference, and Why?* Paper presented at the European Consortium for Political Research Joint Sessions of Workshops, Edinburgh, UK.

Lyons, William, and John M. Scheb II. 1999. "Early Voting and the Timing of the Vote: Unanticipated Consequences of Electoral Reform." *State and Local Government Review* 31:147–52.

Madigan, Francis C., James R. Abernathy, Alejandro N. Herrin, and Clarita Tan. 1976. "Purposive Concealment of Death in Household Surveys in Misamis Oriental Province." *Population Studies* 30:295–303.

Mangat, N. S., and Ravindra Singh. 1990. "An Alternative Randomized Response Procedure." *Biometrika* 77:439–42.

McDonald, Michael P. 2003. "On the Over-Report Bias of the National Election Study Turnout Rate." *Political Analysis* 11:180–6.

Musch, Jochen, Arndt Bröder, and Karl Christoph Klauer. 2001. "Improving Survey Research on the World Wide Web Using the Randomized Response Technique." In *Dimensions of Internet Science,* eds. U.-D. Reips and M. Bosnjak, pp. 179–192. Lengerich: Pabst Science Publishers.

Nordlund, Sturla, Ingar Holme, and Steinar Tamsfoss. 1994. "Randomized Response Estimates for the Purchase of Smuggled Liquor in Norway." *Addiction* 89:401–5.

Orwin, Robert G., and Robert F. Boruch. 1982. "RRT Meets RDD: Statistical Strategies for Assuring Response Privacy in Telephone Surveys." *Public Opinion Quarterly* 46:560–71.

Presser, Stanley. 1990. "Can Context Changes Reduce Vote Over-Reporting?" *Public Opinion Quarterly* 54:586–93.

Rider, Rowland V., Paul A. Harper, L. P. Chow, and I-Cheng Chi. 1976. "A Comparison of Four Methods for Determining Prevalence of Induced Abortion, Taiwan, 1970–1971." *American Journal of Epidemiology* 103:37–50.

Scheers, N. J., and Mitchell C. Dayton. 1982. "The Covariate Unrelated Question Randomized Response Model." In *Proceedings of the American Statistical Association, Section on Survey Research Methods* 407–10.

———. 1987. "Improved Estimation of Academic Cheating Behavior Using the Randomized Response Technique." *Research in Higher Education* 26:61–9.

Shimizu, Iris M., and Gordon Scott Bonham. 1978. "Randomized Response Technique in a National Survey." *Journal of the American Statistical Association* 73:35–9.

Shotland, Lance R., and Lynn David Yankovski. 1982. "The Randomized Response Method: A Valid and Ethical Indicator of the Truth in Reactive Situations." *Personality and Social Psychology Bulletin* 8:174–9.

Soeken, Karen L., and George B. Macready. 1986. "Application of Setwise Randomized Response Procedures for Surveying Multiple Sensitive Attributes." *Psychological Bulletin* 99:289–95.

Stem, Donald E. Jr., and R. Kirk Steinhorst. 1984. "Telephone Interview and Mail Questionnaire Applications of the Randomized Response Model." *Journal of the American Statistical Association* 79:555–64.

Tamhane, Ajit C. 1981. "Randomized Response Techniques for Multiple Sensitive Attributes." *Journal of the American Statistical Association* 76:916–23.

Tezcan, Sabahat, and Abdel R. Omran. 1981. "Prevalence and Reporting of Induced Abortion in Turkey: Two Survey Techniques." *Studies in Family Planning* 12:262–71.

van der Heijden, Peter G. M., Ger van Gils, Jan Bouts, and Joop J. Hox. 2000. "A Comparison of Randomized Response, CASI and Face-to-Face Direct Questioning: Eliciting Sensitive Information in the Context of Welfare and Unemployment Benefit." *Sociological Methods and Research* 28:505–37.

Volicer, Beverly J., Mary H. Cahill, Evelyn Neuberger, and Gretchen Arnt. 1983. "Randomized Response Estimates of Problem Use of Alcohol Among Employed Females." *Alcoholism: Clinical and Experimental Research* 7:321–6.

Volicer, Beverly J., and L. Volicer. 1982. "Randomized Response Technique for Estimating Alcohol Use and Non-Compliance in Hypertensives." *Journal of Studies on Alcohol* 43:739–50.

Warner, Stanley L. 1965. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." *Journal of the American Statistical Association* 60:63–9.

Weissman, Arlene N., Robert A. Steer, and Douglas S. Lipton. 1986. "Estimating Illicit Drug Use through Telephone Interviews and the Randomized Response Technique." *Drug and Alcohol Dependency* 18:225–33.

Williams, Bryan L., and Hoi Suen. 1994. "A Methodological Comparison of Survey Techniques in Obtaining Self-Reports of Condom-Related Behaviors." *Psychological Reports* 7:1531–7.

Wimbush, Dan C., and Donald R. Dalton. 1997. "Base Rate for Employee Theft: Convergence of Multiple Methods." *Journal of Applied Psychology* 82:756–63.

Wiseman, Frederick, Mark Moriarty, and Marianne Schafer. 1975. "Estimating Public Opinion with the Randomized Response Model." *Public Opinion Quarterly* 39:507–13.

Zdep, S. M., Isabelle N. Rhodes, R. M. Schwarz, and Mary J. Kilkenny. 1979. "The Validity of the Randomized Response Technique." *Public Opinion Quarterly* 43:544–9.