

The Validity of the Randomized Response Method in Tax Ethics Research

Ernest R. Larkins, (elarkins@gsu.edu), Georgia State University
Evelyn C. Hume, (ehume@jaguar1.usouthal.edu), University of South Alabama
Bikramjit S. Garcha, (bgarcha@gsu.edu), Georgia State University

Abstract

The survey questionnaire often is used to obtain data in research involving tax compliance and tax ethics issues. The use of the randomized response (RR) survey technique has been suggested for mitigating response and non-response biases in studies regarding sensitive issues. This study compares tax preparer responses using a traditional questionnaire to responses using the RR method. The results indicate that RR does not reduce response or non-response biases in the given decision context. Thus, tax preparers may not be good subjects for a randomized response methodology.

Introduction

Approximately 50 percent of all federal income tax returns are completed by paid tax preparers (IRS 1986). Some evidence suggests that paid preparers may inappropriately reduce the tax liability of their clients at times (Bureau of National Affairs 1986; Podolin 1988; Jackson et al. 1988, 334; Helleloid 1989; Davis 1984). Oftentimes, an unwarranted reduction in tax liability is the result of client pressure (Finn et al. 1988, 607-609). As a result, tax preparers and their impact on tax compliance have received closer scrutiny in recent years (IRS 1988). Research regarding compliance issues can have far-reaching policy implications.

Researchers frequently use a survey questionnaire to obtain data in research involving tax compliance and tax ethics issues.

Comments concerning this article should be sent directly to the authors via e-mail.

Typically, individuals are asked to respond to questions about their tax return filing behavior, or they may be given hypothetical scenarios relating to tax matters and asked to indicate a decision (e.g., see Sanders and Wyndelts 1989; Reckers et al. 1991; Kaplan et al. 1988; Duncan et al. 1989). The individuals are generally assured anonymity and confidentiality with the expectation that this assurance will encourage them to respond completely and truthfully.

A recurring criticism of survey research is that the respondents' self-reported behaviors or decisions do not necessarily reflect their actual behaviors or decisions (e.g., see Akers et al. 1983; Begin and Boivin 1980; Buchman and Tracy 1982; Fidler and Kleinknecht 1977; Hosseini and Armacost 1990). When presented with sensitive questions in a survey, some individuals may not respond truthfully, or they may omit the sensitive questions or decline to participate in the study. Failure to obtain honest responses from the participants introduces a re-

response bias into the data. Failure to elicit responses to some or all of the questions in the survey may result in a non-response bias. In either case, the conclusions drawn from the biased research data may be misleading.

By their very nature, questions asked in tax compliance and tax ethics research are often sensitive; accordingly, truthful responses to such questions are potentially embarrassing and may be incriminating. To ameliorate the problems associated with response and non-response biases in tax research, the use of the randomized response (RR) technique has been recommended (Roth et al. 1989, 224-225; Harwood et al. 1993, 87-88). The RR technique was introduced by Warner (1965) to overcome some of the problems associated with gathering sensitive data. This survey method allows the respondent to answer sensitive questions truthfully without revealing embarrassing or incriminating behavior. Thus, the technique assures a considerable degree of protection in many contexts (Soeken and Macready 1982).

This research compares the results of a direct query survey with the responses of an RR survey asking identical questions to determine whether the latter instrument reduces response and non-response biases in a tax ethics context. The results indicate that the RR technique does not reduce these biases. Since RR procedures are more costly and time consuming than traditional survey procedures, we conclude that an RR methodology should not be used in the decision context of the present study but may be appropriate under other research designs.

Randomized Response Method

On an RR questionnaire instrument, each sensitive question is paired with an alternative innocuous or non-sensitive question. Before responding to each pair of questions, the participant engages in a randomization process such as rolling a die. Depending on the outcome of the randomization process, the respondent answers either the sensitive question or the innocuous question. For example, two questions might be

as follows: (1) Have you ever knowingly failed to report taxable income you received? (2) Were you born in an even-numbered year? The participant might be instructed to roll a die and to answer question one if the outcome is a 5 or 6 and otherwise to answer question two. Several variations of the RR methodology are reviewed in Fox and Tracy (1986).

The researcher does not observe the randomization process and, therefore, does not know its outcome. Therefore, the researcher has no way of knowing whether a participant's response is to the sensitive or to the innocuous question. Even if the respondent can be identified, individual responses remain protected. Theoretically, this protection increases an individual's willingness to participate and, once participating, to provide truthful answers to sensitive questions (Tracy and Fox 1981). Even though the researcher does not know which question was answered by each participant, an unbiased mean and variance for the aggregate pool of respondents can be derived.

The randomization device introduces random error, which reduces the power of statistical testing (Fox and Tracy 1986, 49). To obtain the same power vis-à-vis the traditional questionnaire instrument, therefore, additional responses must be obtained when using an RR methodology. This, in turn, means that use of RR questions is a more costly procedure (e.g., more questionnaires must be mailed). Further, analysis of RR data is more complex than the analysis of traditional survey data. Thus, an RR methodology should be used only if it reduces response bias, non-response bias, or both. Prior studies suggest that RR methodologies have been more successful reducing response bias rather than non-response bias (Boruch and Cecil 1979, 146).

Prior Validation Studies

Validation of the RR technique has been attempted in several social science contexts. For example, the RR method of response elicitation outperformed direct query in studies on illicit

drug use in the military (Reaser et al. 1975, cited in Boruch and Cecil 1979, 71), DUI arrests (Locander et al. 1976), physical punishment of children (Zdep and Rhodes 1976), academic performance (Lamb and Stem, Jr. 1978), public health issues (Greenberg et al. 1970), and criminal arrests (Tracy and Fox 1981). In contrast, the RR technique did not reduce response bias in a survey dealing with multiple topics including racial prejudice, politics, and sexual behavior (Wiseman et al. 1975-76). Note that none of these validity studies involved questions of tax compliance.

Aitken and Bonneville (1980) published research (but not a validity study) that employs an RR methodology in a tax compliance context. Their results were compared with the results of a later survey of taxpayers by Yankelovich, Skelly and White, Inc. (1984, 29), which did not use an RR technique. Though no statistical tests were conducted, the descriptive data suggest that the number of respondents who admitted to cheating on their taxes was significantly greater when the RR method was employed. Specifically, eleven percent of participants admitted overstating tax deductions in the RR survey versus only four percent in the direct query survey. Similarly, 21 percent of the participants in the RR study admitted understating income as compared to only 13 percent in the direct query instrument.

Though the comparison of the two studies above is interesting, there were other differences in addition to the survey technique employed that might explain the differences in noncompliance rates. Among other things, the questions were phrased differently in the two studies, and the studies were conducted in different time periods. The validity of the RR technique in tax compliance research for various populations and various aspects of tax compliance, thus, has not been established.

Whether the RR yields more unbiased results than the direct query method for a population of paid tax preparers, for example, is not yet known. It is possible that tax preparers might feel some affinity toward other preparers (e.g.,

as fellow practitioners or professionals) and, therefore, might decline to answer sensitive questions (resulting in non-response bias) or answer them untruthfully (resulting in response bias). Tax preparers are unlike other survey groups used in validation studies (e.g., drug users) because of this potential affinity. Thus, validation in a tax context is a necessary step to encourage (or discourage) use of the RR methodology in future compliance and ethics studies.

Methodology

This study examines compliance of CPAs with the eight Statements on Responsibilities in Tax Practice or SRTPs (AICPA 1991). Though educational and advisory in nature, the SRTPs provide essential guidance to the CPA in tax practice regarding appropriate standards of responsibilities. There are no separate enforceable penalties or consequences for violating one of the Statements. Nonetheless, they do provide the CPA with a model for ethical decision-making in tax matters.

The SRTPs relate generally to decisions regarding aggressive or questionable tax positions and responses to inaccurate reporting. Briefly, the statements used in this study provide the following: (1) No tax return should be signed if it includes a position that does not have a realistic possibility of being upheld in court or in an administrative hearing; (2) Estimates should not be shown on a tax return in a manner that suggests they are not estimates; (3) Clients should be notified of significant errors on a prior year return regardless of who prepared the return; (4) CPAs should not prepare tax returns that are beyond their capabilities; and (5) Reasonable inquiries should be made when tax return information appears to be incorrect.

The objective of this research is to validate the RR technique in a tax return preparation setting. Specifically, the intent was to determine if the RR method reduces response bias and non-response bias compared with a direct query method in surveys of ethical decisions that CPAs make. No prior studies have compared the ef-

fectiveness of the RR method to that of the direct query questionnaire using professional tax preparers as subjects. Even the comparison in Yankelovich, Skelly and White, Inc. (1984), dealing with taxpayers, was secondary, informal, and noted more as an afterthought. Thus, there have been no formal validations of the RR methodology in any tax compliance contexts. As noted above, tax preparers as a group may differ significantly from other survey groups in which RR has been validated.

To determine whether the RR technique is superior to the direct query method of response elicitation, a direct query survey was mailed to a random sample of 400 CPAs. The sensitive questions on the direct query instrument were identical to those on a separate RR questionnaire. The RR questionnaire was modeled after Berry et al. (1987) and was mailed to 600 CPAs. An example from the RR questionnaire used in this study is presented in Figure 1.

The mailing list for this study was obtained from the California Department of Consumer Affairs. The total design method (TDM) was used in both surveys (i.e., direct query and RR).

Among other things, TDM involves sending a post card reminder four days after the initial questionnaire and sending a follow-up letter with a new questionnaire two weeks later. All envelopes were typed (i.e., peel-off labels were not used), and postage stamps were applied (i.e., instead of franking). TDM has been found to increase response rates in survey studies.

The survey data obtained by direct query were compared to the corresponding information obtained by the RR method to test the following research hypotheses:

HA₁: When facing ethical decisions concerning a client's income tax return, the proportion of CPAs reporting decisions consistent with the Statement of Responsibilities in Tax Practice is greater for those receiving the direct query questionnaire than for those receiving the RR questionnaire.

HA₂: The overall response rate is less for the CPAs receiving the direct query questionnaire than for the CPAs receiving the RR questionnaire.

Figure 1
Example of Question in Randomized Response Questionnaire

You assured your client early in the engagement that a particular expenditure would be deductible on his income tax return. After preparing the client's return but before mailing it, you discover that a Revenue Ruling clearly disallows a deduction for the expenditure in question. After some further investigation, you believe that the Revenue Ruling is a correct interpretation of the law and that there is no reasonable basis for taking the deduction. Nonetheless, you do not think the item will be detected by the IRS. You have basically two choices: (1) You could ignore the Revenue Ruling and mail the return. (2) You could notify the client of the mistake and correct the return.

If the LAST digit of your bill's serial number is 4, 5, or 6, answer question 1A. Otherwise, answer question 1B.

- | | |
|--|-----------|
| 1A. Is the LAST digit of your bill's serial number odd? | Yes _____ |
| | No _____ |
| 1B. Would you notify the client of the mistake and correct the return? | Yes _____ |
| | No _____ |

The first hypothesis tests whether the RR approach reduces response bias while the second tests whether the RR methodology reduces non-response bias.

Analysis and Discussion

The responses to the direct query questionnaire were compared with the responses to the RR questionnaire to check the validity of the RR technique in a tax ethics context. Since the actual instances of unethical decisions is not known, a necessary assumption with this type of validity test is that higher instances of unethical decisions reflect more honest responses (Roth et al. 1989, 226). Thus, RR results that indicate CPAs make unethical decisions significantly more times than do direct query results are evidence that the RR can be used to obtain more reliable data in a tax ethics study.

The mean and variance used to analyze the RR responses appear in Fox and Tracy (1986, 19) and are presented as follows:

$$\mu = \frac{\lambda + P - 1}{2p - 1} \text{ where } p \neq \frac{1}{2}$$

$$\text{Var}(\mu) = \frac{\mu(1 - \mu)}{n} + \frac{p(1 - p)}{n(2p - 1)^2}$$

where:

μ = the estimated probability that the sensitive attribute is possessed; λ = the total proportion of affirmative responses (regardless of whether the response was to the sensitive or the innocuous question); p = the probability that the randomization device would direct the respondent to answer the sensitive question; and n = the total number of responses.

Debriefing questions were included in the questionnaire instrument to eliminate responses of individuals who did not feel protected by the RR procedure. In particular, the responses of those preparers who indicated they disagreed or strongly disagreed that the RR device kept any-

one from knowing which question they answered of each pair presented were excluded from the analysis.

The results from testing the first hypothesis, dealing with response bias, are presented in Table 1. In no instance out of seven did the RR methodology indicate that CPAs engage in unethical practices significantly more often than the direct query indicated. Therefore, the research hypothesis that the RR approach reduces response bias is not supported.

The reason the null hypothesis was not rejected may be that CPAs are hesitant to respond in a manner that might reflect unfavorably on their profession, even with the RR procedure. In other words, they may not be as concerned about the privacy of their own responses (which the RR is said to protect) as they are about the impact their responses might have on the reputation of their profession and their colleagues. This outcome suggests that an RR methodology may not be as appropriate in a tax compliance or ethics study among tax preparers as it would be in some other contexts (e.g., dealing with child abuse or sexual behavior).

The results also do not support the second hypothesis, dealing with non-response bias. Thirty-nine percent of CPAs receiving an RR survey completed and returned the instrument. In contrast, 45 percent of CPAs receiving a traditional direct query instrument responded. Two factors may have contributed to this outcome. First, some CPAs may not have responded to the RR questionnaire because, as Fox and Tracy suggest (1986, 27), they lacked a randomization device (i.e., a dollar bill) at the time. Second, the RR questionnaire required more time and effort to complete than the direct query instrument. CPAs often work under stressful time budgets that provide little free time.

Summary, Limitations, and Conclusion

Two groups of randomly-selected CPAs were surveyed. One group received an RR questionnaire while the other received a tradi-

Table 1
Response Bias Reduction of Randomized Response Versus Direct Query Data

| Question in Survey | Percent of Respondents Not Following SRTPs | | Statistic | P-Values |
|-----------------------|--|--------------|-----------|----------|
| | Randomized Response | Direct Query | | |
| 1 | 0.80% | .00% | 0.14 | 0.557 |
| 2 | 28.84% | 36.52% | -0.95 | 0.172 |
| 3 | -0.53%* | 4.35% | -0.84 | 0.200 |
| 4 | 12.82% | 12.17% | 0.09 | 0.537 |
| 5 | 3.77% | 11.30% | -1.15 | 0.124 |
| 6 | -4.53%* | 3.48% | -1.47 | 0.070 |
| 7 | 0.80% | 4.35% | -0.60 | 0.274 |


*Negative probabilities sometime result when using a randomized response procedure because some participants answer the sensitive question even when directed to answer the innocuous one (Greenberg et al. 1971, 245).

tional direct query questionnaire. Each questionnaire posed seven identical questions dealing with whether the CPA followed the SRTPs. A comparison of the results from each questionnaire revealed that the RR instrument reduced neither response nor non-response bias. Thus, we conclude that a RR methodology is inappropriate when surveying CPAs about their compliance with the SRTPs.

This conclusion may not be generalizable to all CPAs since the sampling frame includes only those practicing in California. The results also do not suggest that RR is inappropriate in other tax compliance or tax ethics contexts. In surveys of taxpayers, for example, the desire to protect the reputation of a profession would be lacking. Supposedly, the feeling of affinity with other taxpayers also would not be as strong as it might among CPAs as a profession. Similarly, non-CPAs working independently or for commercial tax preparation firms might feel less of an obligation to protect the reputation of other non-CPAs who prepare tax returns. Finally, other decisions (e.g., whether to comply with specific provisions of the Internal Revenue Code) may be more sensitive than the decision of whether to comply with the SRTPs, which are only advisory. Thus, the RR technique might provide more reliable data than a direct query

approach in more sensitive contexts. Thus, the results are limited to the specific context.

Suggestions for Future Research

Additional validity studies are necessary for examining other contexts in which the RR research design is appropriate for studying tax compliance and ethics issues. Although taxpayers and tax preparers may be faced with sensitive questions in tax compliance research, the two groups may not be equally reluctant in responding to such questions. Research into the effectiveness of the RR method for different circumstances, environments, and subject groups would allow the RR method to be used in those situations in which the method would be most effective. 

The authors gratefully acknowledge the computer programming assistance of Govind Iyer and the financial support of the College of Business Administration's Research Program at Georgia State University.

References

1. Aitken, S. S., and L. Bonneville, *A General Taxpayer Opinion Survey*, Prepared

- for Office of Planning and Research, Internal Revenue Service, Washington: CSR, Inc., 1980.
2. Akers, R. L., J. Massey, W. Clarke, and R. M. Lauer, "Are Self-reports of Adolescent Deviance Valid? Biochemical Measures, Randomized Response, and the Bogus Pipeline in Smoking Behavior," *Social Forces*, Vol. 62, No. 9, pp. 234-251, 1983.
 3. American Institute of Certified Public Accountants, Tax Executive Committee, *Statements of Responsibilities in Tax Practice*, 1991.
 4. Begin, G., and M. Boivin, "Comparison of Data Gathered on Sensitive Questions via Direct Questionnaire, Randomized Response Technique, and a Projective Method," *Psychological Reports*, Vol. 47, No. 12, pp. 743-750, 1980.
 5. Beldt, S. F., W. W. Daniel, and B. S. Garcha, "The Takahasi-sakasegawa Randomized Response Technique: a Field Test," *Sociological Methods & Research*, Vol. 11, No. 8, pp. 101-111, 1982.
 6. Berry, L. B., G. B. Harwood, and J. L. Katz, "Performance of Auditing Procedures by Governmental Auditors: Some Preliminary Evidence," *Accounting Review*, Vol. 62, No. 1, pp. 14-28, 1987.
 7. Boruch, R. F., and J. S. Cecil, *Assuring Confidentiality of Social Research Data*, University of Pennsylvania Press, Philadelphia, 1979.
 8. Buchman, T. A., and J. A. Tracy, "Obtaining Responses to Sensitive Questions: Conventional Questionnaire Versus Randomized Response Technique," *Journal of Accounting Research*, Vol. 20, No. 2, pp. 263-271, 1982.
 9. Bureau of National Affairs, "Tax Preparers: IRS' Director of Practice Clashes with CPAs over Proposed Ethics Standard," *BNA Daily Tax Reports*, No. 233, pp. G-2, G-3, 1986.
 10. Davis, R. R., "Ethical Behavior Reexamined," *CPA Journal*, Vol. 54, No. 12, pp. 32-36, 1984.
 11. Duncan, W., D. LaRue, and P. Reckers, "An Empirical Examination of the Influence of Selected Economic and Non-economic Variables on Decision Making by Tax Professionals," *Advances in Taxation*, Vol. 2, pp. 91-106, 1989.
 12. Fidler, D. S., and R. E. Kleinknecht, "Randomized Response Versus Direct Questioning: Two Data Collection Methods for Sensitive Information," *Psychological Bulletin*, Vol. 84, No. 9, pp. 1045-1049, 1977.
 13. Finn, D. W., L. B. Chonko, and S. D. Hunt, "Ethical Problems in Public Accounting: the View from the Top," *Journal of Business Ethics*, Vol. 7, No. 8, pp. 605-615, 1988.
 14. Fox, J. A., and P. E. Tracy, *Randomized Response: A Method for Sensitive Surveys*, SAGE Publications, Inc., Beverly Hills, California, 1986.
 15. Greenberg, B. G., J. R. Abernathy, and D. G. Horvitz, "A New Survey Technique and its Application in the Field of Public Health," *Milbank Memorial Fund Quarterly* Vol. 48, No. 10, pp. 39-55, 1970.
 16. -----, R. R. Kuebler, Jr., J. R. Abernathy, and D. G. Horvitz, "Application of the Randomized Technique in Obtaining Quantitative Data," *Journal of the American Statistical Association*, Vol. 66, No. 6, pp. 243-250, 1971.
 17. Harwood, G. B., E. R. Larkins, and J. Martinez-Vazquez, "Using a Randomized Response Methodology to Collect Data for Tax Compliance Research," *Journal of the American Taxation Association*, Vol. 15, No. 3, pp. 79-92, 1993.
 18. Helleloid, R., "Ambiguity and the Evaluation of Client Documentation by Tax Professionals," *Journal of the American Taxation Association*, Vol. 11, No. 3, pp. 22-26, 1989.
 19. Hosseini, J. C., and R. L. Armacost, "Randomized Responses: a Better Way to Obtain Sensitive Information," *Business Horizons*, Vol. 33, No. 3, pp. 82-86, 1990.
 20. Internal Revenue Service, *The Role of Tax Practitioners in the Tax System*, Novem-

- ber, Washington: Office of the Assistant Commissioner, (Planning, Finance and Research), 1988.
21. -----, *Statistics of Income Bulletin* 6, (Summer): No. 8, 1986.
 22. -----, *Survey of Tax Practitioners and Advisers: Summary of Results by Occupation* (June). Washington: Office of Assistant Commissioner (Planning, Finance, and Research), Research Division, 1987.
 23. Jackson, B., V. Milliron, and D. Toy, "Tax Practitioners and the Government," *Tax Notes*, Vol. 41, No. 10, pp. 333-341, 1988.
 24. Kaplan, S. E., P. M. J. Reckers, S. West, and J. Boyd, "An Examination of Tax Reporting Recommendations of Professional Tax Preparers," *Journal of Economic Psychology*, Vol. 9, No. 12, pp. 427-443, 1988.
 25. Lamb, C. W., and D. E. Stem, Jr. 1978. "An Empirical Validation of the Randomized Response Technique," *Journal of Marketing Research*, Vol. 15, No. 11, pp. 616-621, 1988.
 26. Locander, W. S., S. Sudman, and N. M. Bradburn, "An Investigation of Interview Methods, Threat, and Response Distortion," *Journal of the American Statistical Association*, Vol. 71, No. 6, pp. 269-275, 1976.
 27. Podolin, L. "Treasury Raises the Stakes in Circular 230 Proposal," *Journal of Accountancy*, Vol. 169, No. 4, pp. 60-68, 1988.
 28. Reckers, P. M. J., D. L. Sanders, and R. W. Wyndelts, "An Empirical Investigation of Factors Influencing Tax Practitioner Compliance," *Journal of the American Taxation Association*, Vol. 13, No. 4, pp. 30-46, 1991.
 29. Roth, J. A., J. T. Scholz, and A. D. Witte, eds., *Taxpayer Compliance Volume 1: An Agenda for Research*. Philadelphia: University of Pennsylvania Press, 1989.
 30. Sanders, D., and R. Wyndelts, "An Examination of Tax Practitioners' Decisions under Uncertainty," *Advances in Taxation*, Vol. 2, pp. 41-72, 1989.
 31. Soeken, K. L., and G. B. Macready, "Respondents' Perceived Protection When Using Randomized Response," *Psychological Bulletin*, Vol. 92, No. 9, pp. 487-489, 1982.
 32. Tracy, P. E., and J. A. Fox, "The Validity of Randomized Response for Sensitive Measurements," *American Sociological Review*, Vol. 46, No. 4, pp. 187-200, 1981.
 33. Warner, S. L., "Randomized Response: a Survey Technique for Eliminating Evasive Answer Bias," *Journal of the American Statistical Association*, Vol. 60, No. 3, pp. 63-69, 1965.
 34. Wiseman, F., M. Moriarity, and M. Schafer, "Estimating Public Opinion with the Randomized Response Model," *Public Opinion Quarterly*, Vol. 39, No. 1, pp. 507-513, 1975.
 35. Yankelovich, Skelly and White, Inc., *Taxpayers Attitudes Study: Final Report*, Prepared for the Internal Revenue Service, Washington: Government Printing Office, 1984.
 36. Zdep, S. M., and I. N. Rhodes, "Making the Randomized Response Technique Work," *Public Opinion Quarterly*, Vol. 40, No. 1, pp. 531-537, 1976.