Mail Survey Return Rates Published in Health Education Journals: An Issue of External Validity

James H. Price, Judy Murnan, Joseph A. Dake, Jaime Dimmig, and Mary Hayes

ABSTRACT

This study assessed mail survey return rates published in seven general health education journals for the 13-year period, 1990–2002: American Journal of Health Behavior, American Journal of Health Education, American Journal of Health Promotion, Health Education & Behavior, Health Education Research, Journal of American College Health, *and the* Journal of School Health. *A significant difference in mail survey return rates across the seven journals was found. Also, published mail survey return rates significantly increased from 1990–1995 (*M=*61.8%) to 1997–2002 (*M=*65.5%). All of the journals had published a noteworthy percentage (10–26%) of their mailed survey research studies with return rates of less than 50%. Finally, there was not a significant association between sample size and return rates of published mail survey studies. Researchers reporting mail survey research results in health education journals should expect to have return rates of 60% or greater. Yet, such return rates may still be considered a significant threat to the external validity of the findings.*

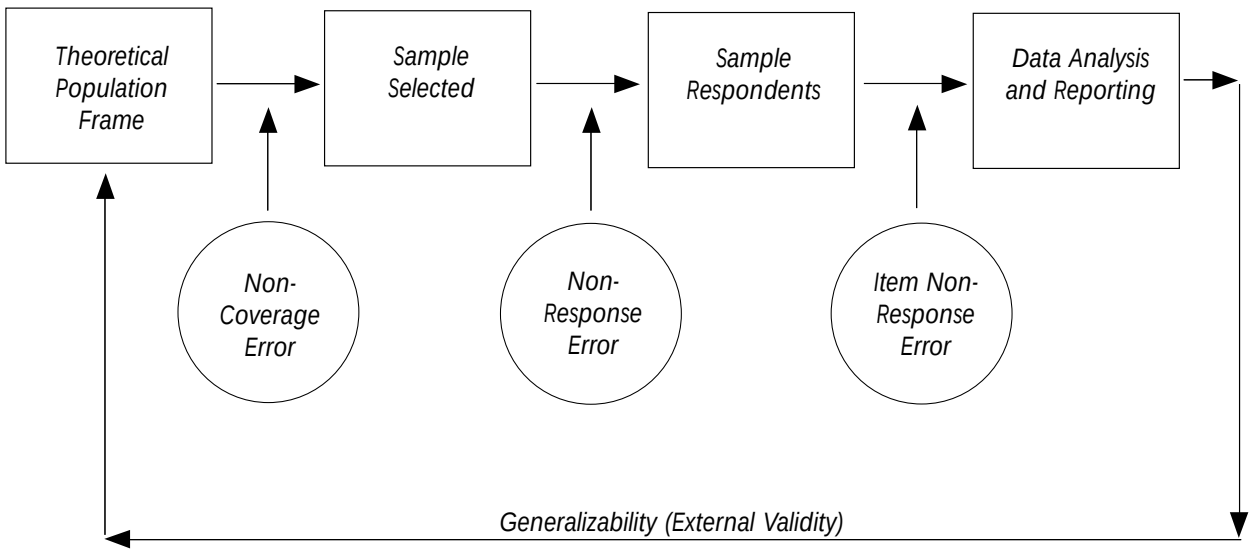
Survey research has many advantages, including convenience, efficiency, low costs, and adaptability. However, the competitive environment of all forms of junk mail attempting to get the attention of potential respondents makes quality survey research more difficult. Fortunately, the scientific base of survey methodology has expanded dramatically (p.6) over the past two decades (Dillman, 2000). A variety of techniques have been studied to significantly improve return rates (Edwards et al., 2002; King, Pealer, & Bernard, 2001).

One of the major methodological issues associated with survey research is external validity. External validity of survey results is concerned with the generalizability of the responses of the sample respondents to the population from which the sample was drawn. When researchers seek to maximize the external validity of survey responses, the researchers need to minimize the potential threats to representativeness. The first potential threat to external validity is how representative the sample selected is of the population (Figure 1). Sampling error can be reduced by use of an appropriate sample selection technique (e.g., random), making sure that the sampling frame does not omit members of the population, and by having an adequate sample size. Another form of this source of error would be when the respondents are not reached after being selected as members of the sample. Thus, they never have the opportunity to respond. The second potential threat to external validity of the survey is nonresponse error. The closer the return rate is to 100%, the greater the likelihood the responses will be representative of the population, assuming the sampling procedure resulted in a truly representative sample of the population. The issue is that those who do not respond to the survey may in some way be different from those who respond to the survey. The nonrespondents may be different from the respondents in that they may hold different beliefs regarding the topic of the survey or they may differ in their demographic characteristics. The third source of error is item nonresponse error. This would include people who respond to the survey but systematically do not answer all of the

James H. Price, PhD, MPH, is a professor of health education in the Department of Public Health at the University of Toledo, Toledo, OH 43606; E-mail: jprice@utnet.utoledo.edu. Judy Murnan, MPH; Jaime Dimmig, MPH; and Mary Hayes are graduate assistants in the Department of Public Health at the University of Toledo. Joseph A. Dake, MPH, PhD, is an assistant professor of health education in the Division of Health at Wayne State University in Detroit, MI 48202.

The authors' recommendation represents a personal preference and is not the policy of the American Journal of Health Education.

Figure 1. Sources of Errors That Create a Potential Threat to the External Validity of Survey Results



questions (e.g., respondents refuse to provide certain answers or fail to respond to a particular question due to questionnaire wording or formatting).

A wide variety of studies have explored nonresponse bias to mailed health surveys. These studies have found that older persons (Criqui, Barrett-Conner, & Austin, 1978; Macera, Jackson, Davis, Kronenfeld, & Blair, 1990), women (Criqui et al., 1978; Bakke et al., 1990; Jooste, Yach, Steenkamp, Botha, & Rossouw, 1990; Paganini-Hill, Hsu, Chao, & Ross, 1993), individuals from higher socioeconomic groups (Bergstrand, Vedin, Wilhelmsson, & Wilhelmsen, 1983; Forthofer, 1983; Sonne-Holm, Sorenson, Jensen, & Schnohr, 1989), persons with higher education (Bakke et al., 1990), those with better health status (Paganini-Hill et al., 1993), and those with more positive health-related behaviors (Bakke et al., 1990; Oakes, Friedman, & Seltzer, 1973; Paganini-Hill et al., 1993) are more likely to respond to health-related mail surveys. However, it would be inappropriate to conclude that representativeness increases monotonically with increasing return rate. As an example, Visser, Krosnick, Marquette, and Curtin (1996) compared the accuracy of telephone surveys with self-administered mail surveys forecasting the outcomes of state elections over a 15-year period. The telephone surveys had return rates of about 60%, whereas the mail surveys had return rates of about 20%. However, the mail surveys predicted election outcomes more accurately than the telephone surveys (1.6% average error vs. 5.2%, respectively). The mail surveys also documented voter demographic characteristics more accurately than the telephone surveys.

The question of representativeness regarding survey return rates is: how much less than 100% are health educators willing to accept for mail survey return rates and still believe the results are representative of the population? The purposes of this study were to answer the following questions: What is a publishable mail survey return rate in health education studies? Do publishable mail survey return rates vary by journal? Has the publishable mail survey return rate changed over time? Are health education journals publishing mail surveys with low (<50%) return rates? Is there a relationship between sample size and return rate for published mailed surveys?

METHODS

Instrument

All published survey studies meeting the following criteria were included in this study: surveys mailed out and returned by mail; surveys distributed in a group setting but mailed back; or surveys mailed to subjects who could then return them to a central location. Excluded studies that had reported return rates included the following examples: classroom distributed and collected surveys; clinic population distributed and collected surveys; workplace group distributed and collected surveys; subjects approached and asked to participate in a program; and telephone or face-to-face interviews. A simple scoring sheet was developed for use with the journals. The data extracted from each mail survey included the following: page numbers of the article, reported survey return rate, size of the respondent sample, or notation that no survey return rate was reported. Survey return rates were checked to make sure the calculation was equal to the ratio of the number of surveys returned divided by the numbers of deliverable surveys distributed. If the numerator or the denominator was absent, then the published return rate was used.

Journals

Only health education journals of a general nature, not area-specific (*Journal of Drug Education, Journal of Cancer Education*) were reviewed for this study. Furthermore, only general health education

Table 1. Mail Survey Return Rates for Seven Health Education Journals (1990–2002)

Journal (Former Title)	Return Rate[A] Lowest	Highest	Mean	(SD)	Less Than 50%
American Journal of Health Behavior (Health Values)	18.0	94.9	59.3	(20.1)	26.4
American Journal of Health Education (Journal of Health Education)	19.9	100.0	62.9	(17.7)	17.9
American Journal of Health Promotion	19.2	95.0	62.7	(16.0)	17.5
Health Education & Behavior (Health Education Quarterly)	25.0	94.0	64.3	(15.4)	13.5
Health Education Research	32.0	98.0	71.6	(15.2)	9.9
Journal of American College Health	12.0	90.0	56.2	(17.8)	13.5
Journal of School Health	28.0	91.0	63.5	(15.4)	15.4

[A]Percentages

journals that were commonly available in academic libraries (at least 25% of the libraries) were included (Laflin, Horowitz, Nims, & Morrell, 2000). The sample of health education journals used for this mail survey return rate analysis included the following seven journals: *American Journal of Health Behavior* (formerly *Health Values*), *American Journal of Health Education* (formerly *Journal of Health Education*), *American Journal of Health Promotion, Health Education and Behavior* (formerly *Health Education Quarterly*), *Health Education Research, Journal of American College Health,* and the *Journal of School Health*. The years 1990 through 2002 were used for the return rate analysis, a 13-year span.

Analysis

The means, standard deviations, and ranges of return rates for each journal were determined. An analysis of variance (ANOVA) was calculated to determine whether significant differences existed in return rates across the seven journals. Tukey *t*-tests were calculated to assess statistically significant post-hoc differences. A *t*-test was calculated to assess differences in return rates based on a median split of the time frame for the years examined (1990–1995 vs. 1997–2002). The year 1996, the median year, was not used in this analysis.

To assess accuracy of identifying reported survey return rates, a sample of three different journals was used and a kappa coefficient was calculated to assess interrater reliability among the five journal reviewers. The kappa coefficient was used to compensate for chance agreement. The mean kappa coefficient was 0.875.

RESULTS

There was a total of 521 published return rates to mailed surveys in the seven health education journals. An ANOVA was calculated for return rates by the seven journals, and it was found that there was a statistically significant difference (F=6.64; df=6; p<.001). Subsequent post-hoc Tukey t-tests found that the statistically significant differences were as follows. *Health Education Research* [mean (M)=71.6%, standard deviation (SD)=15.2] had significantly higher return rates than the following journals: *American Journal of Health Behavior* (M=59.3%, SD=20.1), *American Journal of Health Education* (M=62.9%, SD=17.7), *American Journal of Health Promotion* (M=62.7%, SD=16), *Journal of American College Health* (M=56.2%, SD=17.8), and *Journal of School Health* (M=63.5%, SD=15.4) (Table 1).

A second analysis examined whether there was a significant change over time using a median split for the time period (1990–1995 vs. 1997–2002) in published survey return rates. A *t*-test (t=-2.34; df=482; p=.02) indicated that the published mail survey return rates significantly increased over the two time periods (1990–1995; M=61.8%, SD=18.3 vs. 1997–2002; M=65.5%, SD=16.6).

An analysis assessed how common it was to find survey return rates lower than 50%. A frequency analysis indicated that all of the journals had published a noteworthy percentage (10–26%) of their mailed survey research studies with return rates of less than 50% (Table 1).

A Pearson product moment correlation coefficient was calculated to assess whether a relationship existed between sample size and return rate for 1990–2002. This was done to determine whether studies with larger samples were more likely than those

with smaller samples to be accepted for publication with lower return rates. The result (r=.057 p=.209) was found to be not significant.

DISCUSSION

Surveys that are mailed are useful for examining knowledge, attitudes, and behaviors of subjects who may be geographically widely dispersed, and can be done with the consumption of limited expenditures. However, for mailed surveys to be useful they need higher return rates to reduce the possibility of nonresponse bias. Only one study could be found that examined journal return rates in a particular field of health: medical journals (Asch, Jedrziewski, & Christakis, 1997). This study found that the mean survey return rate of physicians was 54% and for nonphysicians was 68%. The current study found that the mean published return rate for mail surveys for the past 5 years in health education journals was 66%, a number similar to the previously mentioned nonphysician rate.

It was disconcerting to find a sizable percentage of the published mail survey research studies with return rates of less than 50%. Such studies potentially depreciate the quality of published health education research and the journals in which they are found. This study represents the first published study of the health education literature in this area. Hopefully, these results can be utilized to help train survey researchers in health education/health promotion.

The generalizability (external validity) of survey results is critical to the usefulness of survey findings to other professionals. Health educators must understand the potential threats to the external validity of survey results and how to systematically minimize those threats. The first source of error, an inadequate sampling frame, can be minimized by making sure the accessible population is representative of the theoretical population and that probability sampling is used when possible (e.g., simple random sample, systematic sampling with a random start, stratified random sample, or cluster sampling). The second source of error, the one on which this study focused, is a lack of responses to a mailed survey. This can be minimized through a variety of techniques including the quality of the cover letter, attributes of the questionnaire, the return envelope, incentives, and the number of contacts with the potential respondents (King et al., 2001). The third source of error, item nonresponse, can be minimized by using a closed format questionnaire, using vocabulary characteristic of the population, pilot testing the questionnaire, and assessing the questionnaire for stability reliability (test–retest), all of which can facilitate understanding of the questionnaire items (King et al., 2001).

Finally, the delimitations of this study should be noted. First, this analysis included only seven health education journals. Thus, the results of this study may not be applicable to other health education journals (e.g., *American Journal of Health Studies*, *Health Educator*). Second, the results of this study indicated the published return rates for mail surveys only and should not be assumed to be representative of other survey techniques (e.g., interviews, telephone surveys).

CONCLUSION

We urge the editors of health education journals to accept for publication only mailed survey research manuscripts with return rates of 60% or greater. We have selected 60% as the publishable return rate cut-off because of the following rationale. Bias (B) in survey responses is equal to the prevalence (P) of nonresponse times the quantity of response average (R) minus nonresponse average (N); thus B=P x (R-N). Such theory permits survey researchers to estimate (impute) the missing data using responses from multiple wave survey responses (Drane, Rainey, Valois, & Guevera, 1998; Drane, Richter, & Stoskopf, 1993). For example, with a response rate of 60% the bias in the nonresponse rate would be limited. At a nonresponse rate of 40% and a difference of 10% between respondents and nonrespondents, the bias would be 4% (B=0.40 x 0.10=4%). Even a greater difference in responses between respondents and nonrespondents of 20% would cause only an 8% bias in the survey results (B=0.40 x 0.20=8%). If survey researchers can impute the missing data or show that the survey respondents are not statistically significantly different from the population, the nonrespondents, or the chosen sample on key background and demographic characteristics, then and only then should a response rate of less than 60% be published. Additionally, our data indicates that in the early 1990s the average return rate was 62%. Thus, as survey techniques have improved, we should be able to expect to approach the average as the basis for acceptance today. This would be a significant step forward in improving the quality of published survey research. We also strongly encourage that this study be repeated with telephone survey response rates. Such information can be useful to both new (e.g., graduate students) and experienced researchers. Finally, an investigation into how editors/manuscript reviewers select manuscripts for publication that have less than 50% return rates would help shed light on the perceptions of journal editors/manuscript reviewers regarding quality survey research methods.

ACKNOWLEDGMENT

Review of this manuscript was coordinated by a member of the Board of Associate Editors.

REFERENCES

Asch, D. A., Jedrziewski, M. K., & Christakis, N. A. (1997). Response rates to mail surveys published in medical journals. *Journal of Clinical Epidemiology, 50*, 1129–1136.

Bakke, P., Gulsvik, A., Lilleng, P., Overa, O., Hanoa, R., & Eide, G.E. (1990). Postal survey on airborne occupational exposure and respiratory disorders in Norway: Causes and consequences of nonresponses. *Journal of Epidemiology and Community Health, 44*, 316–320.

Bergstrand, R., Vedin, A., Wilhelmsson, C., & Wilhelmsen, L. (1983). Bias due to non-participation and heterogeneous sub-groups in population surveys. *Journal of Chronic Diseases, 36*, 725–728.

Criqui, M. H., Barrett-Connor, E., & Austin, M. (1978). Differences between respondents and nonrespondents in a population-based cardiovascular disease study. *American Journal of Epidemiology, 108*, 367–372.

Dillman, D. A. (2000). Mail and Internet surveys: The tailored design method. New York: Wiley & Sons.

Drane, J. W., Richter, D., & Stoskopf, C. (1993). Improved imputation on nonresponses to mailback questionnaires. *Statistics in Medicine, 12*, 283–288.

Drane, J.W., Rainey, C.T., Valois, R.F., Guevara, E. (1998). Improve mailback questionarre data via a postcard and imputations. *American Journal of Health Studies, 14*(3), 161-168.

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: Systematic review. *British Medical Journal*, *324*, 1183–1191.

Forthofer, R. N. (1983). Investigation of nonresponse bias in NHANES II. *American Journal of Epidemiology*, *117*, 507–515.

Jooste, P. L., Yach, D., Steenkamp, H. J., Botha, J. L., & Rossouw, J. E. (1990). Drop-out and newcomer bias in a community cardiovascular follow-up study. *International Journal of Epidemiology, 19*, 284–289.

King, K., Pealer, L., & Bernard, A. (2001). Increasing response rates to mail questionnaires: A review of inducement strategies. *American Journal of Health Education*, *32*, 4–15.

Laflin, M. T., Horowitz, S. M., Nims, J. K., & Morrell, L. J. (2000). Availability of health education journals in academic libraries. *American Journal of Health Behavior*, *24*, 193–200.

Macera, C. A., Jackson, K. L., Davis, D. R., Kronenfeld, J. J., & Blair, S. N. (1990). Patterns of nonresponse to a mail survey. *Journal of Clinical Epidemiology*, *43*, 1427–1430.

Oakes, T. W., Friedman, G. R., & Seltzer, C. C. (1973). Mail survey response by health status of smokers, nonsmokers, and ex-smokers. *American Journal of Epidemiology, 98*, 50–55.

Paganini-Hill, A., Hsu, G., Chao, A., & Ross, R. K. (1993). Comparison of early and late respondents to a postal health survey questionnaire. *Epidemiology, 4*, 375–379.

Sonne-Holm, S., Sorensen, T. I. A., Jensen, G., & Schnohr, P. (1989). Influence of fatness, intelligence, and sociodemographic factors in a health survey. *Journal of Epidemiology and Community Health, 43*, 369–374.

Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election forecasting? An evaluation of the Columbus Dispatch poll. *Public Opinion Quarterly, 60*, 187–227.