

OMB Section B Statistical Methodology for EIVRI Study

B. Collection of Information Employing Statistical Methods If statistical methods will not be used to select respondents and item 17 on Form 83-I is checked “No” use this section to describe data collection procedures.

1. Respondent Universe and Sampling Methods

The universe from which the study participants will be drawn will be the email list of NIH grant recipients held by the Office of Research Integrity. This list comprises the universe of study subjects, i.e. all approximately 40,000 individuals in any field of study who have received NIH research funds in the past 10 years. This list consists of Research Integrity Officers (RIOs), instructors, lecturers, research administrators, principal investigators (PIs) and others who have requested funding from NIH. For the purposes of this evaluation study, this group is the respondent universe.

We plan to send an email request to participate to all individuals on this list. Some individuals will have moved or retired, some will have invalid email addresses. Moreover, we do not anticipate the response rate to be particularly high. However, by inviting all 40,000 individuals, we anticipate receiving sufficient subgroups, (e.g. by gender, race/ethnicity and age), so that we will be able to conduct subgroup analyses. No other similar studies have been conducted to investigate the effectiveness of the EIVRI to the ORI's knowledge.

2. Procedures for the Collection of Information

A questionnaire has been developed which will collect data on two areas: 1. Demographics of the respondent and 2. Their satisfaction and/or intent to use the DVD in future training sessions. This questionnaire will be developed using Survey Monkey and sent to all invitees, as described in section 1 above. Since the response rate for web surveys can be quite low, all recipients will be sent an email describing the survey and encouraging them to respond in 1-2 weeks. The lowest expected response rate is around 10%, which would result in a phase I response of 4000 cases. While we hope for a better response rate than 10%, we plan to subsample the phase I non-respondents within 1-2 weeks of our initial e-mail message.

Using Survey Monkey as a tool for data collection has several advantages over other types of data collection methods:

1. It enables the survey designer to create complex surveys, which include skip patterns, required answers, multiple or single check-offs and open-ended questions in text boxes;
2. It allows the designer to enter email addresses and keeps track of who responds and who doesn't; this ensures that only non-respondents are re-contacted for follow-up;
3. It allows for data to be downloaded to statistical packages, such as SPSS, for data cleaning and analysis.

The study will use an alpha of 5% for the test accuracy and a regression formula to compare the results.

3. Methods to Maximize Response Rates and Deal with Non-response

Our goal for this web based survey evaluation is to discover whether: 1. Respondents are satisfied with the content of the DVD and 2. They will use the DVD in their research ethics or research integrity classes.

To address nonresponse bias, we will “sample” from this pool of participants in three tiers. Our first request for participation will be sent via email to all individuals on the mailing list. These individuals will then have one week to respond by completing the online survey. At the end of this one-week period, a systematic sample of non-respondents will be taken and these individuals will receive a follow-up email requesting their participation in the study. The individuals contacted as part of this second sampling tier will have one week to respond by completing the online survey. At the end of this one-week period, if the response rate is sufficient, no subsampling will be required. The timing and the size of the subsample of phase I non-respondents will depend on the number of respondents to phase I, the number that complete the entire questionnaire, and the rate that grant recipients e-mail the completed questionnaires.

We need a sample that provides adequate precision, but we don't need responses from all 40000 for that precision. Once we receive responses from approximately 10% of the individuals pooled in phase I, we will have a sample of 4,000 cases and sufficient precision (a standard error of about 1% on estimating a population proportion of .5); however, an estimate based on a 10% response rate may reflect significant nonresponse bias. Therefore, when we reach 4000 responses (or perhaps, earlier if the responses begin to ‘trickle’ in), we will subsample approximately 2,000 of the remaining 36,000 cases to reduce possible effects of nonresponse bias. Depending whether the combined weighted response rate of the phase I and phase II subsample reaches 70-75% or not, we may attempt a 3rd phase email request to some or all of the phase II non-respondents. We will halt all follow-up activities once we reach an adequate response rate or when we determine that further follow-up would not be productive. At each stage of follow-up after the first phase mailing, we will modify the email message by reminding the non-respondents that their response is very important to represent the opinions of the entire universe of those pooled.

The participants who decline to participate after reading the informed consent statement at the beginning of the survey will be asked to provide a reason as to why they do not wish to take part in the study. This will aid in analysis of nonresponse.

4. Tests of Procedures or Methods to be Undertaken

Pilot test: The outcome measures for this study will be pilot tested before being used in data collection. We have requested nine interested parties to pre-test the evaluation questions (for the pilot study). From their responses, we will obtain feedback on the overall design of the questionnaire, as well as potential questions that should be included in the questionnaire. Please see a copy of the questionnaire (see Appendix A). The data used for this study will be collected from electronic questionnaires. The independent variables are the subject demographics. The dependent variables will include the satisfaction and use questions. We will use both qualitative and quantitative analyses for this study. The proposed quantitative analysis will consist of univariate descriptive statistics of the demographic variables and the individual items for each of the satisfaction and use questions. Cross tabulations will be done with key demographic, satisfaction, and use variables. The open-ended questions will allow some qualitative analysis. The analyses of this data will also be included in the final report.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Statistician:

Douglas Wright, MS, DSFederal Inc.