

SUPPORTING JUSTIFICATION – PART B

Information and Communications Technology (IC) Needs Assessment OMB Control Number 2130-New

B. Collections of Information Employing Statistical Methods

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

Respondent Universe

This needs assessment is focused on the active labor segment of the railroad industry because it is the largest portion of the industry and the target of most industry educational outreach and knowledge dissemination efforts. We would like to fill in the gaps in understanding across this segment.

It is not realistic to survey the entire labor population for several reasons. The total size of the railroader population is unknown. Further, even if we could identify every railroader, the effort required to perform a census far surpasses current budget constraints. While convenience sampling would likely be more cost effective, that method would not allow us to infer results to the rest of the industry because we would not be sure the sample would be representative of the industry (Crano & Brewer, 2002; Isaac & Michael, 1997). It is more practical to obtain a random sample of a portion of the industry and generalize their responses to the rest of that portion of the railroad industry.

For this research, we intend to target the International Association of Sheet Metal, Air, Rail and Transportation Workers-Transportation Division (SMART-TD) and the Brotherhood of Locomotive Engineers and Trainmen (BLET). These two unions provide an almost complete coverage of the industry, from engineers and conductors to yard and signalmen. Each of these unions represents broad job types so that adequate coverage of the industry will be provided from within their ranks.

(Note: Some conversations have occurred with BLET and SMART-TD representatives about this survey. There has been positive response to those conversations, but buy-in from decision makers has not yet been obtained. During the preliminary public response phase, these commitments will be obtained.)

Given the two unions' website data and recent studies by the FRA, we estimate that 180,000 railroad train and engine (TY&E) labor, non-retired, workers comprise the combination of freight and passenger rail, making the potential respondent universe 180,000 TY&E service employees in the United States. Most U.S. TY&E employees -- 125,000 -- are members of the SMART Union. The remainder -- 55,000 -- are members of the BLET. This estimate was arrived at by Internet research and the researchers' understanding of the industry and is used to define

sample size and stratification across the unions. Small adjustments (+/- 20,000) will have minor effects on sample size needs. These two groups provide a near comprehensive universe of U.S. TY&E employees. A very small number of TY&E employees work for short line railroads and are not represented by a labor organization. As such, they will not be part of this study.

Both labor organizations maintain a database with the names, mailing addresses, date of birth, and date joined the union for all its members. **Both the BLET and SMART-TD (will) have agreed to generate a random sample of their members in accordance with the method described herein.** The database will be current as of the date that the random sample is drawn. Only actively working (i.e., non-retiree) TY&E employees living in the United States will be included in the sampling frame. Full-time union officials and non-TY&E workers who are union members will be excluded from the sampling frame. Respondents will be selected randomly from this sampling frame. The only type of stratified random sampling from the union databases that is possible is a geographic stratification. However, this study does not seek to look at the characteristics of subpopulations defined by geography. If job type were available in the union database, then a stratified sample on this basis would be used.

2. Describe the procedures for the collection of information including:

- Statistical methodology for stratification and sample selection,
- Estimation procedure,
- Degree of accuracy needed for the purpose described in the justification,
- Unusual problems requiring specialized sampling procedures, and
- Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

Sample size needed

Confidence in generalizations is set to the industry standard of 95% and a 50/50 distribution of proportions is assumed because it is the most conservative and this study is not focused on testing any particular population differences.

Sampling error is directly related to acceptable margin of error and for this study we are setting this at +/- 5%. This choice is based on the balanced consideration of effort and value. Reducing the margin of error may add to the precision of our understanding of this population, but in turn would significantly increase the number of responses necessary to generalize with some validity (e.g., at 3% the required sample would go from 383 to 1,056). With the selection of a 5% margin of error, if we find that 59% of our sample strongly agrees with a statement we can say with some confidence that the population’s response would be between 54% and 64%. Given the descriptive nature of this study, this margin of error is acceptable.

This formula is used to compute sample size (Dillman, 2007, p.206):

$$n = \frac{(N)(p)(1-p)}{(N - 1)(B/C)^2 + (p)(1-p)} \quad \text{Equation (1)}$$

- n = sample size needed
- N = population size
- P = proportion of population differences

B = acceptable amount of sampling error
C = Z statistic associated with the confidence level; 1.96 for 95% level

Using an estimated 180,000 labor railroaders with a 5% margin of error, the following computation (Equation 2) estimates a sample size of 383.

$$n = \frac{180,000 * .5 * (1-.5)}{(180,000 - 1) (.05/1.96)^2 + (.5) * (1-.5)} \quad \text{Equation (2)}$$

3. *Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.*

Response Rate

This collection has not been conducted previously, so the response rate must be estimated from other similar efforts and research on survey methods. In previously approved studies with subgroups of the TY&E employee population, the response rates ranged from 21-50%¹. Most of those studies included substantial incentives and required more extensive participation: \$5 with the invitation and a \$75 gift card for returned surveys and sleep logs. While one study, with a similar population, reported a 36% response rate with no incentives, we found no documentation of the methods that were used to garner that rate.

This study does not intend to offer such a large incentive (e.g., a \$75 gift card) to encourage participation and cannot find otherwise adequate documentation to support making reliable assumptions about an expected response rate for this population. Therefore, we use existing research outside the railroad community to inform our response rate calculation. Dillman (2007) has found that by applying some incentive and outreach efforts surveys can obtain 20-30% response rates. Therefore, and with no evidence to support the assumption that the railroad industry participants will be more or less likely to participate, we will oversample with the assumption of a 25% response rate. Attaining approximately 383 respondents will allow our results to be generalized to the population with a 5% margin of error. Therefore, an initial sample size of 1,532 (1,532*.25=383) is recommended.

Of the 180,000 members between the two unions, the SMART-TD represents 62.5% of that population and the BLET represents 37.5%. Therefore, the stratum defined for each union is designed to represent those percentages. The total number of potential participants is rounded up to 1,533.

¹ OMB 2130-0558/0570/0577/0588 Sleep Studies garnered 49.9%/46%/33%/21% response rates. A Track Inspection Study (oddly also numbered 2130-0588) garnered 36% response rate with no incentives.

Table 1. Sampling stratification for TY&E workers in the BLET and SMART-TD unions

Union	Sampling Frame	% of Frame	Stratum N	Total
SMART-TD	125,000	62.5%	.625 x 1,532	958
BLET	55,000	37.5%	.375% x 1,532	575
Total	180,000	100%	100%	1,533

Both labor organizations maintain databases with the names, mailing addresses, date of birth, and date joined the union for all of its members. The unions will generate a random sample in accordance with the above table for our use that includes these variables:

- First and last name
- Mailing address
- Birthdate
- Year joined the Union

Efforts to Encourage Participation

An announcement letter on official FRA letterhead will be sent out in advance of the survey. This announcement is intended to raise the awareness of the forthcoming survey and explain the importance of their participation (Dillman, 2007, 2011). We are recommending that the announcement letter include union and industry leaders’ signatures to convey the importance of the coming invitation.

One week after the announcement letter is sent, the survey with a cover letter will be sent. Postcard reminders will be sent two weeks after that and again two weeks later, to all non-respondents reminding them of the survey and its importance. Two weeks later, six weeks after the initial mailing, a final mailing will go out to those who have not yet completed the survey. This final mailing will include another copy of the survey to encourage their completion.

- The announcement letter will be from an FRA Administrator. In addition to explaining the procedure, this letter will describe the benefits of the study and encourage individuals to participate. This letter, as well as an article we will write for inclusion in various train newsletters prior to the study, are intended to familiarize members with the effort and to legitimize the study.
- The survey materials will be sent via first class mail with a personally addressed and dated letter printed on high quality paper. The package will contain a cover letter, the questionnaire, and a first-class postage paid envelope for return of the questionnaire.
- At 2-week intervals, two follow-up reminder postcards will be sent to all recipients who have not returned their questionnaire and who have not indicated that they do not wish to participate.
- The questionnaire is 8.5 x 11 in., printed on white paper with no questions on the

cover page. The cover page contains only the title “Information and Communications Technology Questionnaire,” the FRA form number, the OMB control number, and the participant’s ID number. The required OMB statement, including the assigned OMB Control number and the confidentiality assurance statement, is on the inside of the cover page of the questionnaire. The questionnaire instrument will be visually pleasing and easy to read.

Analysis of Possible Non-Response Bias

If the response rate for any key item falls below 70 percent, we will conduct a nonresponse bias analysis using standard analysis procedures (i.e., OMB, 2006). If the overall response rate for the study falls below 80 percent, we will conduct a nonresponse bias study. From the union membership databases, it is possible to determine the age and length of union membership of each TY&E employee. The nonresponse bias study will involve comparing the age distribution of the survey non-respondents with the age distribution of the respondents. The mean age for each group will be compared. For several reasons, age is an important characteristic for assessing potential bias. First, research has shown that use of technology is more prevalent with younger users. In addition, age is highly correlated with years of work experience and seniority. If a Z test shows that the respondents and non-respondents have similar age distributions, then we will be confident the respondents reflect the larger TY&E population.

- 4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.*

Pilot Study

Experts in survey development and the railroad industry have reviewed and provided input into the existing instrument. A pilot study will be undertaken to improve item wording. For the pilot study, a convenience sample of 20-30 railroaders, personally known by us, will be contacted and asked to complete the survey. If they agree, the introductory packet described herein will be sent to them. Railroaders will be encouraged to provide input on the survey items on a 1-on-1 basis by calling the indicated point of contact. Concurrently, a stakeholder panel of key industry representatives will be convened to review the instrument. This input will be incorporated into the instrument. Data collected from these opportunities will also be compared to the final respondents’ data to identify possible non-response biases. Based on our experience with the industry, it is expected that those known to us will be more likely to complete the survey and will be more connected to ICT.

Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Data collection and analysis will be led by the Volpe National Transportation Systems Center.

Heidi D. Howarth, Ph.D.

(617) 494-2522

Kim Davies-Schrils

(617) 494-2583

With consultation on statistical aspects and survey design from:

Michael A. Harnar, Ph.D., iBiz/CORA

(909) 524-7800

Juna Z. Snow, Ph.D., iBiz/CORA

(617) 440-4435

References

- Crano, W. D., & Brewer, M. B. (2002). *Principles and Methods of Social Research* (2nd ed.). Mahwah, New Jersey: Lawrence Earlbaum Associates
- Dillman, D. A. (2007). *Mail and Internet Surveys: The tailored design method* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Dillman, D. A. (2011). Reconsidering Mail Survey Methods in an Internet World. American Statistical Association/American Association for Public Opinion Research Webinar, April 13, 2011. Downloaded April 6, 2014 from http://www.umb.edu/editor_uploads/images/u54/ASAAPORwebinar4_11_11.pdf
- Isaac, S., & Michael, W. B. (1995). *Handbook in Research and Evaluation: For Education and the Behavioral Sciences* (3rd ed.). San Diego, CA: Educational and Industrial Testing Services.
- Office of Management and Budget (2006). *Standards and guidelines for statistical surveys, September 2006*.