

Guide for Measuring Compliance Assistance Outcomes

Revised June 2002



Office of Enforcement and
Compliance Assurance (MC 2222 A)

EPA300-B-02-011
June 2002
www.epa.gov/oeca

TABLE OF CONTENTS

	Page
SECTION I: INTRODUCTION	1
A. Background	1
B. History and Purpose of This Guide	1
C. Overview of Guide	2
SECTION II: COMPLIANCE ASSISTANCE ACTIVITIES AND OUTCOMES	4
A. Compliance Assistance Activities	4
B. Measuring Results of Compliance Assistance	5
C. Continuum of Output and Outcome Measures	7
SECTION III: HOW TO PLAN AND DESIGN AN ASSESSMENT	8
A. What Is the Goal of Your Compliance Assistance Project?	8
B. What Is the Purpose and Scope of Your Evaluation?	9
C. What Measures Are Appropriate?	12
D. Which Data Collection Method Best Meets Your Needs?	13
E. Do You Need an ICR?	18
SECTION IV: HOW TO GET THE MOST OUT OF YOUR SURVEY	21
A. Questionnaire Design	21
B. Survey Implementation-The Tailored Design Method	29
C. Where Do We Go From Here?	38
D. Lessons Learned from Measurement Projects	46
SECTION V: AN INTRODUCTION TO STATISTICAL SAMPLING	48
A. Key Terms	49
B. Statistical Validity	50
C. Sampling Frame-The List of Potential Respondents	50
D. Estimating A Sample Size	53
E. Drawing the Sample From the Population	60
SECTION VI: OECA’S EXPERIENCE WITH THE DILLMAN METHOD AND STATISTICALLY VALID SURVEYS	62
A. Background and Purpose of the Study	62
B. Methodology	64
C. Implementation	68
D. Results	69
E. Lessons Learned	86

APPENDIX A:	Performance Profile for EPA’s Enforcement and Compliance Assurance Program
APPENDIX B:	Office of Compliance Guidance on the Need for Information Collection Requests (ICRs) for the Evaluation of Compliance Assistance Activities
APPENDIX C:	How to Obtain Clearance for Regional Compliance Assistance Evaluation Surveys Under the Generic ICR 1860.01 OMB Control #2020.0015
APPENDIX D:	Menu of Sample Survey Questions by Outcome Measure
APPENDIX E:	Examples of Letters Used by OECA to Gather Data by the Dillman Method
APPENDIX F:	Advanced Data Analysis - Measuring Association and Making Comparisons
APPENDIX G:	Sampling From Unknown Populations

LIST OF FIGURES

	Page
2-1 Continuum of Compliance Assistance Measures	7
3-1 Do You Need An ICR?	19
4-1 Example Pie Chart for Frequency Distribution	40
4-2 Example Chart of Frequency Distribution	41
4-3 Example of Answer Distribution Chart	43
5-1 Formulas for Estimating Sample Sizes	59
6-1 Metal Finishing Performance Evaluation Survey	90
6-2 Marina Checklist	95

LIST OF TABLES

	Page
3-1 Outcome Measures	12
3-2 Pilot Project Response Rates	13
3-3 Compliance Assistance Data Collection Tools	16
3-4 OECA's Recommended Follow-up Methods by Assistance Activity	17
4-1 Subset of Dillman's Principles of Questionnaire Construction	28
4-2 Example Frequency Distribution	40
6-1 Background Information on EPA Data Collection Study	63
6-2 Factors Contributing to Sample Size Estimate	64
6-3 Response Rates	71
6-4 Mailed Survey Returns	72
6-5 Distribution of Responses for Site Visits	73
6-6 Cost Per In-Scope Response	74

6-7	General Characteristics of the Two Samples in the Metal Finishing Sector	76
6-8	Types of Compliance Assistance Received For the Two Samples in the Metal Finishing Sector	76
6-9	Awareness of Regulatory Requirements for the Two Samples in the Metal Finishing Sector	77
6-10	Characteristics of Marinas in Both Samples	78
6-11	Compliance Comparison for Metal Finishing Sector	80
6-12	Summary of Compliance Question Results for Metal Finishing – Comparison of Mailed Surveys and On-Site Observation Compliance Rates	81
6-13	Hazardous Waste Compliance for Marinas	82
6-14	Oil and Fuel Compliance for Marinas	83
6-15	Hazardous Materials Compliance for Marinas	84
6-16	Storm Water Compliance for Marinas	84
6-17	Summary of Compliance Question Results for Marinas–Comparison of Mailed Surveys and On-Site Observation Compliance Rates	85

SECTION I: INTRODUCTION

A. Background

Congress enacted the Government Performance and Results Act (GPRA) in 1993 to encourage federal managers to measure success based on the results of their projects instead of simply the level of activity occurring under their programs. The act requires all federal agencies to develop 5-year strategic plans with goals, objectives, and performance measures. In response to GPRA, EPA developed a strategic plan that delineates goals to achieve its objectives. Goal 9 of EPA's strategic plan requires the Agency to ensure full compliance with laws intended to protect human health and the environment. EPA plans to achieve this goal through compliance incentives and assistance programs, identifying and reducing significant noncompliance in high-priority program areas, and maintaining a strong enforcement presence in all regulatory program areas.

The Office of Enforcement and Compliance Assurance (OECA) has developed a set of measures to evaluate its performance toward achieving Goal 9. This set of measures includes outcome measures (changes in behavior due, at least in part, to compliance assurance activities), environmental indicators (measures of progress toward achieving environmental or human health objectives), as well as output measures (measures of the numbers of activities). These measures apply only to EPA's federal enforcement and compliance assurance program. They do not serve as a framework for measuring performance of state enforcement and compliance assurance programs. OECA and the states developed a separate set of accountability measures for state enforcement and compliance assurance programs, incorporated in Performance Partnership Agreements (PPAs). Thus, while OECA's measures are not applicable to states directly, states can still find information in this guide to help measure goals articulated in their PPAs. Appendix A includes a performance profile for EPA's Enforcement and Compliance Assurance program.

B. History and Purpose of This Guide

OECA first published this document in March 1999 to help Regions measure the results of compliance assistance being delivered through a series of pilot projects. Since that time, Regions and states have gained a great deal of experience measuring results. In addition, EPA has received feedback

on the 1999 guide from the Compliance Assistance Advisory Council (CAAC), participants in Regional training programs, and other stakeholders. This revised version of the *Guide for Measuring Compliance Assistance Outcomes* responds to that feedback and incorporates lessons learned from the federal pilot projects and state measurement projects. The purpose of this revision is to provide the community of compliance assistance providers and interested parties with the best possible guidance for measuring their results. This version of the guide also adds information on conducting statistical studies.

OECA believes that measuring results is a key step to making better management decisions and complying with GPRA. Evaluations conducted under this guide are consistent with GPRA requirements¹ and are compatible with the requirements of EPA's Regional Compliance Assistance Tracking System (RCATS). In addition, EPA has incorporated lessons learned from the pilot projects into its strategic plan and annual performance plans. EPA management has also used the results of measurement projects to support internal management and policy decisions.

C. Overview of Guide

This guide consists of the following sections designed to help you measure the outcomes of compliance assistance:

- # **Section II: Compliance Assistance Activities and Outcomes** discusses the types and purposes of compliance assistance activities and the outcomes associated with these activities.
- # **Section III: How to Plan and Design an Assessment** provides a step-by-step overview of key planning and design issues and describes the benefits and limitations of different data collection tools.
- # **Section IV: How To Get The Most Out Of Your Survey** includes tips on writing good surveys and suggestions on how to get more out of your evaluation using the Tailored Design Method.
- # **Section V: An Introduction To Statistical Sampling** provides a simplified discussion of how to conduct a statistically valid evaluation.

¹ Prior to publishing this revised guide, EPA has not used the results of the pilot projects for GPRA purposes due to concerns of the subjectivity of previous surveys. Through this guide and related projects, OECA is putting forth a renewed effort to collect statistically valid data and to improve the quality of data collected through nonstatistical methods.

- # **Section VI: OECA's Experience with the Dillman Tailored Design Method** discusses the recent Metal Finishing and Marina Compliance Assistance Program Evaluation Study.²

² Don Dillman is the Thomas S. Foley Distinguished Professor of Government and Public Policy in the Departments of Sociology and Rural Sociology, and Deputy Director for Research and Development in the Social and Economic Sciences Research Center (SESRC) at Washington State University. The Tailored Design Method is a comprehensive approach to designing and implementing self-administered surveys. A complete exposition of the method can be found in Don Dillman's book *Mail and Internet Surveys: The Tailored Design Method* (John Wiley and Sons Inc., 1999).

SECTION II: COMPLIANCE ASSISTANCE ACTIVITIES AND OUTCOMES

Compliance assistance consists of providing information and technical assistance to the regulated community to help it meet the requirements of environmental law. First and foremost, compliance assistance attempts to ensure that the regulated community understands its obligations by providing clear and consistent descriptions of regulatory requirements. Compliance assistance also can help the regulated community find cost-effective ways to comply and to go "beyond compliance" in improving their environmental performance through the use of pollution prevention and other innovative technologies.

A. Compliance Assistance Activities

EPA groups compliance assistance activities into four major categories: telephone assistance, workshops, presentations, compliance assistance tools, and onsite visits.

- # **Telephone Assistance** includes assistance provided by hotlines, where the telephone is the primary outreach vehicle for compliance assistance.
- # **Workshops** include training sessions and seminars that are sponsored by the compliance assistance program and that involve a group of regulated entities or assistance providers. Workshops are more substantial than presentations and tend to involve greater participant interaction.
- # **Presentations** are specific compliance assistance materials communicated to a group of regulated entities at meetings that may or may not be sponsored by the compliance assistance program. Presentations include speeches, multi-media demonstrations, and panel discussions. Presentations are briefer and less resource intensive than workshops.
- # **Compliance Assistance Tools** include printed materials (e.g., newsletters, fact sheets, information packets, brochures), videos, slide shows, and Web sites that are produced by the Regions, Headquarters, or others for distribution purposes. Examples of compliance assistance tools include plain language guides, self-audit checklists, expert systems, and CD-ROM-based or Web-based applicability determinations.
- # **Onsite Visits** include visits to potentially regulated facilities to provide technical assistance, compliance assistance, environmental management reviews, and pollution prevention assistance. Onsite visits may also be used to set a baseline from which programs can measure the facility's progress. Onsite visits do not include inspections where the intended purpose is to carry out enforcement.

B. Measuring Results of Compliance Assistance

EPA has identified three types of measures to gauge the success of compliance assistance activities. These measures include *output measures*, *outcome measures*, and *environmental and public health impact indicators*. While this document focuses on outcome measures, understanding output measures and environmental and public health impact indicators helps put the outcome measures into perspective.

Output Measures

Output measures are defined as “quantitative or qualitative measures of important activities, work products, or actions taken by EPA or by states under delegated federal programs.”³ They assess both the number of EPA products or activities (e.g., the number of fact sheets developed, the number of onsite visits conducted) and the number of facilities reached through compliance assistance activities (e.g., the number of helpline calls answered, the number of people at a workshop, the number of guides downloaded from the Internet).

Outcome Measures

Outcome measures are “quantitative or qualitative measures of changes in behavior of the public or regulated entities caused, at least in part, by actions of government.”⁴ Outcome measures include changes in understanding⁵, changes in behavior, and site-specific environmental and human health improvements:

- # **Changes in understanding** reflect an increased knowledge of regulatory or nonregulatory environmental issues, including reporting and monitoring requirements,

³ U.S. EPA. 1997. National Performance Measures Strategy (NPMS). <<http://es.epa.gov/oeca/main/fedgov/npms.html>> December. p. 4.

⁴ NPMS, p. 4.

⁵ The original guide combined awareness and understanding into one measure. Feedback from the Regions implementing the pilot projects and CAAC suggested that these two measures be separated as they are distinct conditions. EPA considers changes in understanding more indicative of results than simple changes in awareness. EPA now considers awareness a component of “reach”, an output measure.

regulatory schedules, and pollution prevention opportunities. An example of *changes in understanding* includes measuring the percentage of facilities receiving assistance that indicate an improved understanding of environmental regulations or the number of facilities attending a workshop that gained knowledge about pollution prevention or control technologies. Changes in understanding can most effectively be measured by testing knowledge before and after the workshop.

Behavioral changes represent actual changes that a regulated entity has undertaken as a result of compliance assistance. Examples of *behavioral changes* include the number of facilities that submitted required permit application or notification forms because of a training program, or the number of facilities that adopted recommendations discussed during an onsite visit. *Behavioral changes* can be voluntary (e.g., voluntary implementation of pollution prevention technologies as a result of publication of pollution prevention guidance documents or fact sheets) or regulatory (e.g., facilities reporting overlooked chemicals as a result of the publication of Toxic Release Inventory guidance documents). Improvements in compliance are also included under behavioral change.

EVALUATING BEHAVIORAL CHANGES - THE NATIONAL NITRATE COMPLIANCE INITIATIVE

In 2000, EPA implemented the National Nitrate Compliance Initiative, which educated facilities in the metal finishing sector about Emergency Planning and Community Right-to-Know Act (EPCRA) §313 requirements and EPA's Audit Policy. Through a phone survey, Region 3 found that 78 percent of the respondents (14 facilities) indicated a behavioral change as a result of the mailing. Examples of actions taken included:

- # Conducted an audit;
- # Made a process, operating, or material change;
- # Developed an internal monitoring/reporting system;
- # Researched alternatives and substitutions; and
- # Obtained further technical assistance

Environmental and human health improvements are measures of environmental and human health improvements at specific facilities resulting from compliance assistance activities. Examples of *environmental and human health improvements* would be the number of pounds of pollutant emission reductions at a facility that adopted a control technology explained in a training video, or the number of facilities reducing chemical exposure to workers as a results of practices presented at a workshop.

Environmental and Public Health Impact Indicators

Environmental and public health impact indicators are defined as quantitative or qualitative measures of progress over time toward achieving national environmental or human health objectives.

These indicators help EPA measure what impacts its environmental programs are having on national environmental problems. Environmental indicators might, for example, show a reduced level of nutrients in a water body over a specified amount of time.

C. Continuum of Output and Outcome Measures

Figure 2-1 shows the continuum of output and outcome measures starting from reaching your targeted population to having the population reduce pollution as a result of your efforts. Each measure builds on the previous measure on the continuum. Changes in behavior will not occur until the target audience understands the regulatory requirements. Similarly, it is difficult to assess the site-specific environmental and human health improvements without knowing the changes in behavior.

Although this document focuses on measuring the outcomes (i.e., *changes in understanding, changes in behavior, and environmental and human health improvements*) associated with compliance assistance projects, understanding how effectively you have reached the target audience will help you measure these outcomes. If the hotline, assistance tool, or workshop is reaching only a small portion of the intended audience, there will be limited corresponding changes in understanding and behavior. For example, if only a few printers in a targeted community are aware of a compliance assistance workbook and hotline, only a small number of facilities can possibly make changes as a result of the assistance.

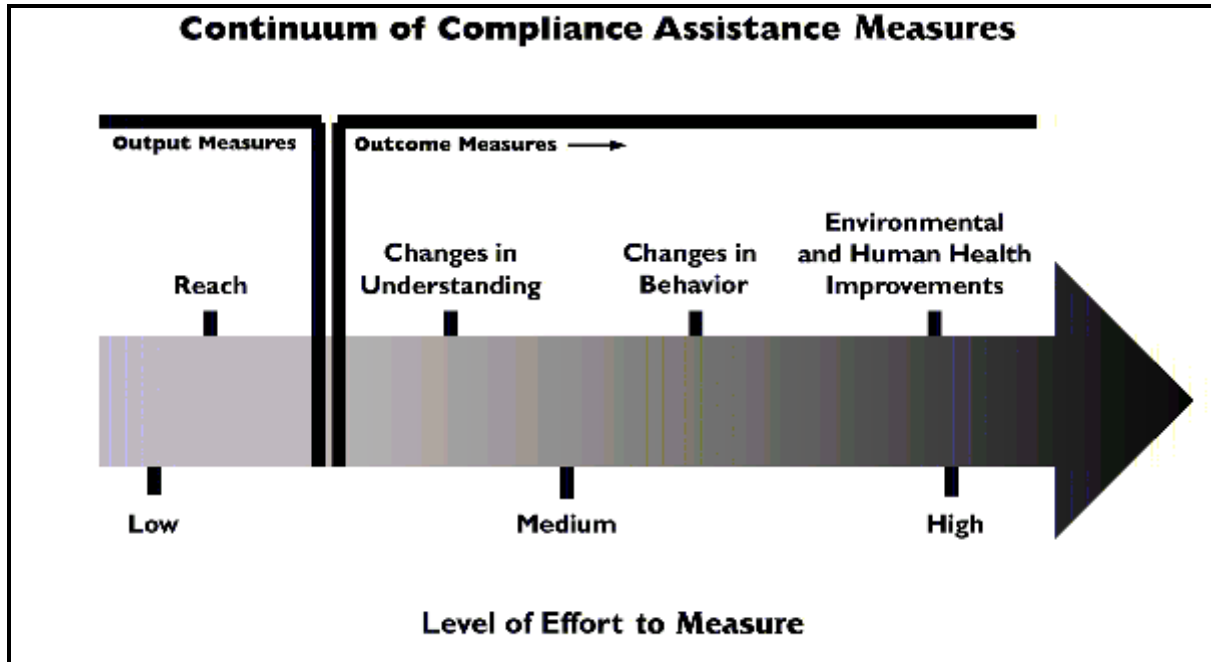


Figure 2-1. Continuum of Compliance Assistance Measures

SECTION III: HOW TO PLAN AND DESIGN AN ASSESSMENT

This section will assist you in planning and designing evaluations of your compliance assistance projects. This is the most critical stage in evaluating your success. Without effective planning and design, the subsequent steps in an evaluation are likely to provide results that are inconclusive and difficult to understand. Identifying the goals of your compliance assistance project and where on the continuum (Figure 2-1) you are starting from, and defining the purpose and scope of the evaluation, will help you determine the best approach for your evaluation. To begin planning your evaluation, answer the following key questions:

- # What is the goal of your compliance assistance project?
- # What is the purpose and scope of your evaluation? (Should you collect statistical or anecdotal data?)
- # What measures are appropriate to determine if you have met your goal?
- # Which data collection method best meets your needs?
- # Does the Paperwork Reduction Act Apply?

A. What Is the Goal of Your Compliance Assistance Project?

The first step in the evaluation process is to identify the goals of your compliance assistance project. Is the goal of your project to achieve environmental results, such as improving the storage and disposal of hazardous waste by regulated facilities? Or, is the goal to increase understanding of reporting requirements? Understanding where your project falls on the compliance assistance continuum will help ensure that you select appropriate measures to evaluate your success. For example, through a 2001 pilot project, Region 7 promoted compliance with Resource Conservation and Recovery Act (RCRA) requirements, including proper storage and disposal methods, to colleges and universities. Since the goal of this project was to motivate behavioral change (e.g., proper storage of waste) and achieve environmental outcomes (e.g., proper disposal of waste), measuring both behavioral change and environmental improvements was appropriate. In another example, Region 5 conducted a pilot project to improve Toxic Release Inventory (TRI) reporting. The Region's goal was to make sure that facilities understood how they should properly report. For this project, measuring changes in understanding TRI reporting requirements and forms before and after the workshop was appropriate.

REGION 9'S GOAL: IMPROVED COMPLIANCE IN OPPRESSING HAZARDOUS AIR POLLUTANTS

In support of one of EPA's national priorities, Region 9 set a goal to improve compliance with the Chrome National Emission Standards for Hazardous Air Pollutants (NESHAP) at chrome plating facilities in the Clark County area of Nevada. The Region provided onsite compliance assistance tailored to address problems identified in earlier visits. During subsequent revisits, inspectors were able to accurately measure the result of the initiative—100 percent compliance with all parts of the NESHAP at every facility.

Examples of specific project goals include:

- # Improve the quality of TRI data by helping the regulated community understand the EPCRA reporting requirements;
- # Raise understanding of RCRA requirements via seminars to help prevent improper hazardous waste management;
- # Reduce the high number of chemical accidents in the ammonia refrigeration industry by providing a sector-specific compliance assistance manual focused on preventing accidental spills; and
- # Assist colleges and universities in meeting their environmental regulatory obligations through forums on regulations that apply to educational institutions.

B. What Is the Purpose and Scope of Your Evaluation?

Carefully assessing the evaluation’s purpose—why the evaluation is needed, how the information will be used, and who will use the evaluation results—will help you determine the scope of the evaluation. A key decision in determining the scope is whether to use a statistically valid study or an anecdotal assessment. For example, if you want to generalize to the overall population or to compare two groups (such as those that received compliance assistance and those that did not), then you might want to consider conducting a statistical study. On the other hand, if you are collecting information such as lessons learned, innovative techniques used by facilities as a result of compliance assistance, or how the compliance assistance activity helped the audience, then an anecdotal assessment will probably meet your evaluation needs. Whichever option you choose, consider using observable data (e.g., review self-reported data, conduct onsite revisits) to document results.

Statistically Valid Evaluations

If you want to generalize to the overall population or to compare two groups (such as those that received compliance assistance and those that did not), you might want to consider conducting a statistical study. Statistical evaluations enable you to generalize your evaluation results to a larger

audience (e.g., all facilities in an industrial sector, geographic region, or all users of a compliance assistance tool). This type of study requires additional up-front planning in identifying and selecting study participants, and may require that you collect data from a larger number of respondents to make reliable generalizations. Statistically valid surveys may require more resources than anecdotal assessments (discussed below). See Section V for guidance on how to conduct statistically valid evaluations.

Consider using a statistically valid assessment if you want to:

- # Generalize results to an overall population
- # Compare two groups
- # Evaluate the overall impact of a program

EVALUATING THE IMPACTS OF COMPLIANCE ASSISTANCE ON METAL FINISHERS

EPA conducted a statistically valid assessment of metal finishers in Regions 1 and 5 to identify areas requiring additional compliance assistance (e.g., RCRA, air regulations TRI reporting) and to evaluate the effectiveness of compliance assistance programs. By surveying a sample of 100 metal finishers, EPA was able to extrapolate results to the two regions.

Anecdotal Assessments

Anecdotal assessments—evaluations that describe accomplishments, yet make no broad generalizations or claims⁶—are suitable for most compliance assistance evaluations conducted by OECA. Anecdotal assessments tell a story about how compliance assistance has impacted the group of people that responded to your survey. These evaluations can provide some quantitative information, such as the number of facilities changing behavior as a result of receiving a compliance assistance tool, and can also ascertain how and why the facilities responded to the compliance assistance. However, results from these assessments cannot be scaled up to a larger group. In general, an anecdotal assessment relies on surveys of parties receiving compliance assistance. One problem of anecdotal assessments is that, unless there is a very high response rate, they are subject to nonresponse bias (i.e., the results of those who responded may be significantly different from those who did not respond). To minimize this potential bias, use the data gathering practices outlined in Section IV, to maximize your response rate. A well-thought-out and well-received anecdotal assessment can provide useful information for making a decision as to whether this activity is worth continuing, and can be worthwhile if sufficient resources are not available for a statistical study.

Consider using an anecdotal assessment if you:

- # Follow methods to get a high response rate
- # Do not expect nonresponse bias
- # Want to receive an indication of whether you have met your goals

⁶ Sparrow, Malcolm K. 1997. Regulatory agencies searching for performance measures that count. *Journal of Public Administration Research and Theory*.

INCREASING UNDERSTANDING OF HAZARDOUS WASTE MANAGEMENT ON THE U.S./MEXICO BORDER

Region 6, in partnership with the Texas Natural Resource Conservation Commission (TNRCC), found that warehouses along the U.S./Mexico border were violating RCRA requirements because of a lack of knowledge of RCRA and hazardous waste management. In response, Region 6 and TNRCC developed a compliance assistance seminar designed to improve understanding of RCRA. Region 6 conducted a survey of the seminar to learn how to improve the compliance assistance and to determine whether it was effective and worth continuing. Results indicated that 86% of respondents would like more similar seminars made available, and 50% made changes in environmental practices or took other action to comply with the regulations as a result of the seminar.

Using Observed Data

Whether you intend to conduct a statistically valid study or an anecdotal assessment, you should consider ways you can directly observe changes made by recipients of your compliance assistance. Direct observation—measuring changes through pre-and post-tests, observing changes onsite, or assessing self-reported documents—tends to be more reliable than surveys that simply ask respondents to indicate whether or not they made changes.

You can use direct observation in a variety of ways. For example, for workshops, you might consider assessing the participants' understanding of regulations before and after the workshop. Region 6 conducted a workshop designed to teach warehouse employees about RCRA requirements and hazardous waste management. Since the goal was to increase understanding, the Region administered pre-and post-tests to determine the increase in knowledge of RCRA requirements (the result was a 29 percent increase in test scores) and receive feedback about the specific strengths and weaknesses of the material presented. Another option might be to visit facilities to see if they have, in fact, made changes to comply with the regulations for which they received assistance. Region 8 used inspections to evaluate the effectiveness of multimedia compliance assistance delivered to auto service and repair shops. During the visits, Region 8 inspectors directly observed and measured increases in compliance and beyond-compliance actions. Finally, you might consider tracking compliance data that your audience is required to report, such as TRI reports, EPCRA §312 reports, and permit applications. For example, if you conduct a workshop on air permitting requirements, check to see how many of the attendees applied for permits 6 months after the workshop. Also, check to see if the quality of the reporting has improved.

C. What Measures Are Appropriate?

Once you establish the purpose and scope of the evaluation, you will need to define what measures you will use to evaluate the success of the program. Defining success up front, based on the project goal and purpose of the evaluation, will help you select the appropriate evaluation measures, as well as interpret the results. Try to identify what results you would expect based on your experience with similar projects and the goals you established at the outset of the activity. Table 3-1 lists specific measures developed by OECA for assessing the outcome of compliance assistance activities; select from this set of measures in developing your project. OECA developed these measures based on its analysis of the results of pilot projects. OECA tracks these measures on a national level through the Regional Compliance Assistance Tracking System (RCATS). Use the descriptions of different measures in Section II to decide which measures best fit your needs and resources.

Table 3-1. Outcome Measures

Outcome Measure	
Changes in Understanding	Number of facilities/respondents who say they better understand how to comply with environmental regulations
Behavioral Changes	
Environmental and Human Health Improvements	Number of facilities that reduced emissions or other pollutants
	Number of facilities indicating a safer environment for workers
	Quantified environmental improvements (e.g., amount of emissions or pollutants reduced)

D. Which Data Collection Method Best Meets Your Needs?

Gathering the necessary data to evaluate compliance assistance activities is usually the most difficult, time-consuming, and resource-intensive step in the evaluation process. It is essential that you select the most appropriate data collection tool to meet the goals of your evaluation.

Data Collection Tools

Whether you conduct an anecdotal assessment or statistical study, there are a number of data collection tools available. These tools include surveys (mail, fax, email, Internet, and phone) and observed data (onsite revisits, pre/post tests, and reviews of self-reported data). Where possible, OECA recommends using observed data as they are more objective than changes reported by survey participants. In general, you can collect more detailed information using telephone surveys and onsite revisits than using mail, fax, email, or online surveys, since they allow you to ask follow-up or clarifying questions. Another consideration in selecting a tool is the expected response rate; Table 3-2 presents response rates for surveys conducted as part of a set of federal pilot projects. OECA does not recommend using mail-back comment cards (tear sheets); EPA and states found that these cards were potentially biased and received very low response rates (usually around 2 percent).

Table 3-2. Pilot Project Response Rates

Measurement Method ¹	Range of Response Rates	Average Response Rate
Pre-/Post-tests (1)	97%	97%
Phone survey (3)	41% - 100%	80%
Mail survey (7)	2% - 46%	17%
Email survey (1)	100%	100%
Workshop survey (1)	85%	85%
Revisits (3)/Inspections (2)	100%	100%
Inspections (2)	100%	100%

¹ The number of projects on which the data are based is given in parentheses.

Additional considerations for each of the data collection tools are discussed below:

- # **Pre-/Post-Test.** Before conducting the compliance assistance activity (e.g., workshops, training sessions), consider testing attendee knowledge of regulations you plan to cover. At the end of the compliance assistance activity, retest the participants to determine changes in understanding of the materials presented. Similarly, you can assess behavioral

practices at the facility before a workshop and practices reported in a follow-up survey to identify changes made. Pre-/post-tests can also help you improve your compliance assistance materials by revealing areas where key messages did not come across. Pre-/post-tests conducted at workshops are considered part of the workshop and are exempt from the Paperwork Reduction Act.

Telephone Survey. A telephone survey is a standard set of questions asked to potential respondents over the telephone. These surveys, used alone or in combination with mail or online surveys allow you to ask follow-up or clarifying questions, potentially resulting in better data than a mailed survey. Telephone surveys work best if the list of potential respondents is a manageable number (OECA recommends less than 50 respondents). To reduce costs, some regions have hired college students to make the call-backs.

Mail/Email/Fax Survey. A mail, email, or fax survey is a set of questions sent to potential respondents with a request that they voluntarily respond. These surveys enable you to reach a large number of potential respondents, and may be the best option where there are more than 50 recipients. However, mail/email/fax surveys can provide ambiguous results, since it is not easy to immediately follow up and clarify unclear, conflicting, or unexpected responses. Similarly, a limited level of detail is obtained, as respondents will generally not spend the time to write long answers to open-ended questions.

Online Survey. An online survey is a set of questions posted on a Web site or list server. These surveys have the potential to reach a large number of respondents. For surveys on Web sites, you can reach users that might otherwise be unknown to you. Many respondents like online surveys because they can respond at their convenience and they do not need to worry about losing a survey or mailing it back. As with mail surveys, however, the online survey may provide only limited detail as respondents might not want to spend time typing in a longer response. In addition, without follow-up, there is potential for ambiguity or conflicting results, as with the mail survey. Another problem with a simple online survey gauging the effectiveness of a Web site is that respondents may be

biased. For example, only respondents who found a Web site to be very useful typically respond to the survey while those who did not find the site useful or found it only marginally useful typically do not respond to the survey. To reduce nonresponse bias you may want to advertise the surveys on relevant list serves to give non-regular user an opportunity to respond.

Onsite Revisit. Onsite revisits involve returning to facilities that previously received a compliance assistance visit. Revisiting facilities can provide excellent data since you can use direct observation to make assessments

**IN DENVER, DIRECT OBSERVATION
EQUALS RELIABLE ASSESSMENTS**

Region 8 delivered multimedia compliance assistance, pollution prevention, and cost savings tools to auto service and repair shops in the Denver metropolitan area. Via onsite revisits, Region 8 inspectors measured behavioral changes (including changes in compliance and beyond-compliance actions), and also looked for changes in awareness and understanding. By using direct observation, the inspectors were able to make reliable assessments and learn about the business (and language) of auto service and repair.

and because facilities are likely to spend the necessary time to answer questions while you are on site. In addition, the revisit itself might spur additional compliance assistance or pollution prevention activities. Revisits can be performed by either compliance assistance staff or by inspectors. For the best results and highest level of cooperation, this data gathering should not be linked to enforcement initiatives. If revisits are performed by inspectors, it is critical that the information collected for the compliance assistance evaluation is not used to make the site a target for enforcement.

- # **Data Reviews** . Reviewing in-house data such as permits, permit applications, and TRI forms can also help you assess changes in understanding and behavior. For example, if your project sought to improve the quality of TRI Form R submissions or stormwater permits, you can assess the quality of these documents before and after the training or compare submissions from those receiving assistance and those who did not receive assistance to gauge your progress.

You can also consider a combination approach, such as using a mail and/or online survey with a phone survey to selected participants or to reach non-respondents. Such an approach enables you to potentially reach a large number of respondents, yet also collect detailed information from selected participants. In addition, it enables you to ask follow-up questions and clarify any unexpected results from the mail survey. In general, the more resources you expend, the more responses and detail you will receive. Table 3-3 highlights the uses of the different tools, the resource considerations, average response rates (for OECA pilot projects), and tips for lowering costs and improving response rates.

Table 3-3. Compliance Assistance Data Collection Tools

Method	When to Use	Resource Considerations	Range of Response Rates ¹	
Comment Cards	OECA strongly discourages using mail-back comment cards (tear sheets). EPA and states found that these cards were potentially biased and received very low response rates (usually around 2 percent).			
Pre-/Post-Tests	<ul style="list-style-type: none"> To measure changes in understanding 	Does not require OMB approval if conducted with workshop. Requires staff time to administer and adds time before and after workshops/presentations.	97%	
Telephone Survey	<ul style="list-style-type: none"> To obtain detailed data. 	Long-distance charges will apply if surveying a large geographical area. Staff time to administer telephone survey is higher than the mail or online survey. Staff will need to be trained to ensure survey consistency.	41-100%	Call respondents either early or late in the day to bypass secretarial screening. Also consider using non-EPA employees to conduct the survey to improve perception of anonymity.
Mail/Email/Fax Survey	<ul style="list-style-type: none"> To reach many respondents. To obtain "yes/no" answers. 	Mailed surveys require postage costs for sender and respondent. Reminder postcards require additional postage. Email and fax surveys can be limited by technical (e.g.; computer hardware) capabilities.	2-46%	
Online Survey	<ul style="list-style-type: none"> To evaluate electronic services. Can also be used as an alternative to a mail survey. 	Up-front costs include posting the survey on the Web site and programming electronic data entry capabilities, if desired. Can eliminate data entry costs associated with other surveys.	N/A - Total population will not be known.	Prompt Web users to respond to the survey by announcing it to the target audience through list servers, newsletters, conferences, etc., and by making it noticeable on the Web site. Online surveys exclude compliance assistance users without access to Internet and can self-select favorable responses.

¹ From OECA pilot projects.

Selecting the Proper Evaluation Method

In general, the evaluation method used should match the type and intensity of compliance assistance activity. It makes more sense, for example, to use onsite revisits to follow up on previous visits than it does to follow up on those who simply received a guidebook or called for telephone assistance. Table 3-4 offers OECA's recommended follow-up methods for each type of compliance assistance activity.

Table 3-4. OECA's Recommended Follow-up Methods by Assistance Activity

Compliance Assistance Activity	Preferred Follow-up Method	
Onsite visits	1. Onsite revisit; or 2. Phone survey	1. Do at least 25% revisits; or 2. Call 100% of sites
Workshop/presentations	1. Pre-/Post-tests; and 2a. Phone survey if <50 attendees; or 2b. Mail survey if >50 attendees	1. Test all participants; and 2a. Phone all participants; or 2b. Use the Tailored Design Method (TDM) ⁷
Compliance guides distributed through workshops	1. Phone survey if <50 attendees; or 2. Mail survey if >50 attendees	1. Phone all participants; or 2. Use the TDM No mail back cards
Compliance guides distributed via mail	Mail survey	1. Maintain a list of mailed recipients and use the TDM No mail back cards
Compliance guides distributed via the Internet	No follow-up for outcome measurement; count number of guides downloaded only	No mail back cards
Hotlines	Periodic user survey	Use phone surveys
Compliance assistance centers - Internet	Online survey	Use a secured site that would eliminate participants from taking the survey multiple times

⁷ See Section IV of this guide for more information on using the Tailored Design Method.

E. Do You Need an ICR?

Congress enacted the Paperwork Reduction Act (PRA) to protect businesses, states, and the general public from burdensome paperwork. The act requires federal agencies⁸ to obtain approval from the Office of Management and Budget (OMB) prior to collecting substantially similar information from more than nine nonfederal respondents. As defined by the PRA, a “collection of information” means obtaining or soliciting information through identical questions, or reporting or recordkeeping requirements. The PRA applies to both mandatory and voluntary data collection efforts; therefore, most compliance assistance evaluations are subject to the PRA. Note, however, that the following actions do *not* require OMB approval:

- # Surveys handed out onsite immediately after a workshop, seminar, or meeting that ask participants about the quality of the seminar (e.g., knowledge of speakers, usefulness of handouts);⁹
- # Pre-/post-tests conducted at a workshop, seminar, or meeting; and
- # Attendance sign-in sheets at a meeting, workshop, or Web site.

Appendix C provides additional information to help you determine whether or not your evaluation falls under the PRA.

Figure 3-1 should help you determine what next steps are necessary to proceed with your evaluation. In general, when you ask the same set of questions (whether voluntary or regulatory) to more than nine people, the survey will fall under the PRA¹⁰. OECA has, however, obtained an expedited approval process for many compliance assistance surveys through its generic ICR. You can use the generic ICR when your goal is to determine the effectiveness of compliance assistance activities on the audience that receives the compliance assistance (e.g., participants at a workshop or users of a compliance assistance tool). Note, however, that the generic ICR *cannot* be used when you plan to use a

⁸ The PRA applies to federal projects only. The act does not apply to states unless they are conducting a survey on behalf of EPA. See Appendix B for more information.

⁹ Surveys to collect baseline data to assess awareness and understanding prior to the workshop *may* require OMB approval. Consult Appendix C for specific scenarios.

¹⁰ Note that this requirement does *not* apply when contacting officials at federal facilities.

statistical approach to generalize the effectiveness of a compliance activity on an overall population. In this case, you will have to develop a separate ICR for your evaluation.¹¹

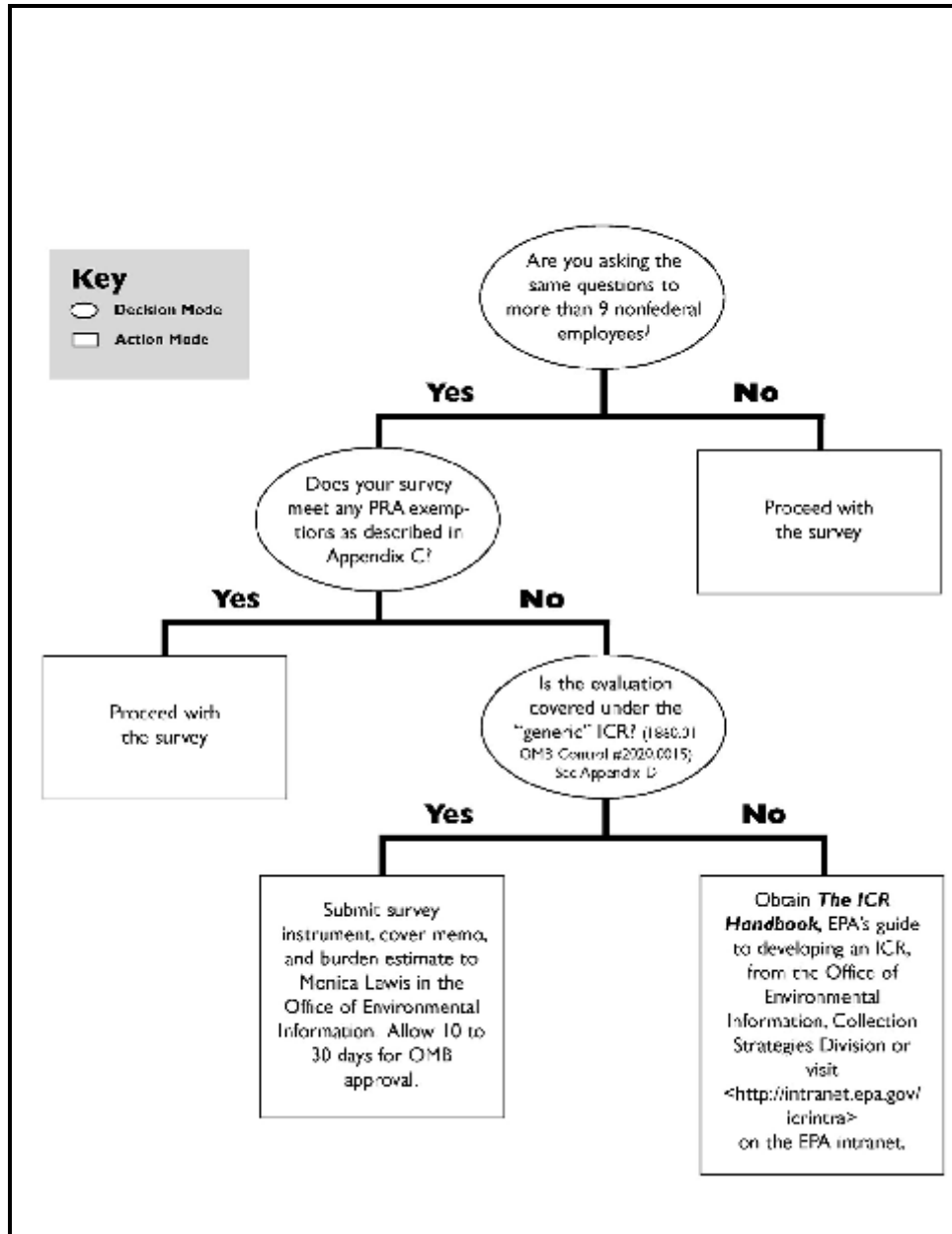


Figure 3-1. Do You Need an ICR?

¹¹ OMB put this constraint on the ICR to give the public an opportunity to comment on survey methodology prior to implementation. For guidance on developing an ICR, see EPA's *ICR Handbook* available from the Office of Environmental Information, Collection Strategies Division (CSD). CSD provides policy direction and oversight of Agency management of regulatory information and manages the Agency's administration of the burden reduction provisions of the Paperwork Reduction Act. CSD's OECA liaison, Monica Lewis, will answer any questions you might have about whether an ICR is needed, issues involved in preparing an ICR, or the ICR clearance process. Monica can be reached by calling 202-566-1678 or via mail at Mail Code 2822T.

If your survey is covered under the generic ICR, you will still need to obtain clearance from OMB before you distribute your survey. To obtain clearance, you need to submit a copy of the survey instrument, cover memo, and burden estimate to Monica Lewis in the Collection Strategies Division (CSD) of the Office of Environmental Information. Appendix B contains the Office of Compliance Guidance on the Need for Information Collection Requests (ICRs) for the Evaluation of Compliance Assistance Activities. Appendix C contains information on how to obtain clearance for regional compliance assistance evaluation surveys under the generic ICR 1860.01, OMB control # 2020.0015. Appendix C can also be found on the Web at <www.epa.gov/oeca/perfmeas/icrfacts.html>.

SECTION IV: HOW TO GET THE MOST OUT OF YOUR SURVEY

This section will help you design clear and effective questionnaires, techniques to implement efficient surveys to provide high response rates, and organize your findings in a meaningful way. OECA developed this guidance by consulting survey experts and analyzing the results of pilot projects. This section draws on Don Dillman's¹² work in designing and implementing surveys. OECA strongly encourages you to follow these suggestions, which can help you maximize the effectiveness of your evaluation efforts.

Section A discusses how to write good questions and how to construct the questionnaire and Section B discusses how to effectively implement the survey. Section C discusses and defines basic data analysis techniques. Section D discusses lessons learned from recent regional projects. Appendix D for samples of survey questions by outcome measure.

A. Questionnaire Design

The questionnaire development phase of your survey project is very important. Numerous evaluation projects have failed to yield reliable information because the survey instrument has been flawed (Dillman, 1999). A poorly constructed questionnaire can lead to increased nonresponse, misinterpretation of questions (resulting in unreliable information), incomplete answers to questions, and unanswered questions.

The most important aspect of questionnaire design is whether the questions are salient to the respondent. That is, the respondent should feel the need to provide the information you are requesting. Salience is created through writing good questions and organizing those questions effectively, as discussed below.

¹² Don Dillman is the Thomas S. Foley Distinguished Professor of Government and Public Policy in the Departments of Sociology and Rural Sociology, and Deputy Director for Research and Development in the Social and Economic Sciences Research Center (SESRC) at Washington State University. He is recognized as one of the leaders in survey research in United States and has over 32 years of experience in designing and conducting survey research. His books are considered seminal texts in how to effectively design survey research.

Tips on Writing Effective Questions*Some Questions You Should Ask About Each Question*

To begin, once you have formulated a set of questions, you should review those questions to determine if they are something that respondents will want to answer and if they will provide the information that you need. Dillman (1999) suggests that you ask the following eight questions about each survey question.

Is the question written so that each respondent can provide an answer? If your questions provide a set of potential answers, make sure that all possible answers are covered. That is, no respondent should look at the set of potential responses and not find an answer that fits their situation.

To what extent do survey recipients already have an accurate, ready-made answer for the question they are being asked? To ensure that the question is answered, you should make sure that you are asking something that respondents can answer with relatively little effort. When possible, you should avoid including questions that require respondents to look up information or to make calculations.

Will the respondent be able to accurately recall past events and behaviors related to the question? You should consider whether or not the information you are requesting is something that the respondent can recall easily.

Will the respondent be willing to reveal this information? You should consider whether or not the respondents will reveal the information you are asking for and how truthful those answers will be. This is especially critical in evaluating information related to compliance with environmental regulations.

Will the respondent feel some motivation to answer the question? Although you may find the answer to a question interesting, you should consider whether or not the respondent would feel the same way. Research has shown that the more salient that a question is to a respondent, the more likely that respondent will answer the question.

Will the way in which the question is presented influence the answers that the respondents may give? You should think about the answer categories that you have provided to the respondent and consider whether or not those will have some influence on the answers provided. The *Principles of Writing Good Questions* section below discusses some of these considerations.

Are the data being collected in more than one manner? Some survey efforts use more than one mode of collecting information (e.g., mail and personal interviews). If you want the data collected to be consistent across the modes, you need to make sure that the mode you use will not

influence the answers you get. The *Principles of Writing Good Questions* section below discusses some of these considerations.

Is it possible to change the wording or structure of a question? Although a question may be poorly written and have a number of faults, it may not be open to change. For example, you may want your data to be consistent with a previous survey effort and therefore you would need to use the same question wording.

Open-Ended Versus Closed-Ended Questions

Questions can be divided into two basic categories: open-ended and closed-ended. For open-ended questions, the surveyor has not provided a set of answers from which the respondent can choose. Closed-ended questions, on the other hand, provide a set of answers from which the respondent can choose. One of these question types is not necessarily better than the other in all survey situations. Nevertheless, you may want to consider writing many of your questions as closed-ended since it is relatively easier to construct reliable closed-ended questions than reliable open-ended ones.

Open-ended questions work well when they have unambiguous interpretation. For example, the question “How old are you?” is an open-ended question if you simply supply a blank space for the answer. You may also want to give the respondent free rein on answering a question. These may be exploratory questions (e.g., “What environmental compliance issues do you see confronting your business in the near future?”) or you may be seeking feedback on a program you have instituted (“How might we improve this seminar?”).

Not all closed-ended questions are easy to construct. Dillman (1999) divides closed-ended questions into “ordered” and “unordered” categories. Ordered closed-ended questions provide the respondent with a set of ordered responses (e.g., like or dislike on a scale of one to five). In an ordered question, the respondent must simply place herself on a scale. Unordered questions provide a set of mutually exclusive categories that have no particular order (e.g., choosing among a set of statements that best describes a respondent opinion or situation). In an unordered question, the answering task is much more complicated since respondents must compare one answer to all other answers. Thus, respondents may find unordered questions with many answer possibilities more difficult to answer than some open-ended questions since more evaluation must be done.

Dillman's Principles for Writing Good Questions

Dillman (1999) provides a set of 19 principles for writing good questions. We have highlighted the key steps. You should refer to Dillman's book¹³ for the complete list and discussion of the rationale behind each principle as well as insightful examples.

1. **Choose simple rather than specialized words.**
2. **Use as few words as possible in posing the question.**
3. **Use complete sentences.**
4. **Avoid vague quantifiers when more precise ones can be used.** For example, use quantifiers such as "1-2 times per month" or "less than ten" rather than "rarely" or "just a few."
5. **Avoid requesting a specific answer to a question if the respondent may not be able to provide such an answer easily.** For example, it may be too specific if you provide a blank space for the question "How many autos were repaired in your shop in 2001?" However, if that same question supplied a set of ranged responses (e.g., "less than 100," "100-200," etc), then the question may be easier for a respondent to answer.
6. **For scaled questions, use an equal number of positive (e.g., agree) and negative (e.g., disagree) categories.** For example, the question "How useful did you find our training seminar?" could be followed by four answer categories: "very helpful", "somewhat helpful", "somewhat unhelpful", and "not helpful at all".
7. **Place "neutral" categories (e.g., "neither agree nor disagree") in the middle of scales, but place "undecided" categories at the end of scales.**
8. **For unordered closed-ended questions, make all comparisons (i.e., answers) equivalent to avoid bias.** For example, the question "What factors have led to environmental issues for your industry?" that provides "Irresponsible chemical management" and "Economic hardship" does not provide equivalent answers. The use of "irresponsible" makes that answer a value judgement while the other answer is more objective.
9. **For agree/disagree type questions, mention both sides in question.** For example, use "To what extent do you *agree or disagree* with the following statements?" rather than "To what extent do you *agree* with the following statements?"

¹³ Dillman, Don, 1999. *Mail and Internet Surveys: The Tailored Design Method*, John Wiley and Sons, Inc: New York.

10. **Avoid “Check All That Apply” questions.** Research into survey design has found that respondents tend to answer check off “just enough” of the response options to form what they consider to be a satisfactory answer, rather than completely evaluating *each* response option.
11. **Be sure that answer categories are mutually exclusive.**
12. **For questions that require the respondent to recall past events, ask a few simple questions regarding the details surrounding the event to improve respondents’ recall.** For example, you may want to know respondents’ thoughts on a seminar. To improve recall, ask some simple questions about the seminar (“Did you attend by yourself or with others?”, “Did you have to travel a long distance to get there?”). Research has shown that these techniques can assist respondents in recalling the event more clearly.
13. **If you are comparing the data you are collecting with other data (e.g., previous survey efforts, census data), be sure that the wording allows that comparison to be made.** In other words, make sure you are asking the same questions as the original survey. This is particularly important for projects where data from two surveys done at different times (i.e., pre and post compliance assistance efforts) will be compared.
14. **Avoid using double negatives.**
15. **Do not use double-barreled questions.** For example, the questions “Did you find the seminar and the materials we passed out afterwards helpful?” asks the respondent to evaluate two things at once (the seminar and the materials) and to provide one answer for both. It could be that some respondents found the seminar helpful, but not the materials.
16. **If you have to ask objectionable questions, find a way to word the question to reduce the impact of that question.** For example, the question “Does your facility have compliance problems?” can be interpreted as threatening, but “Has your facility ever had trouble in complying with some regulatory requirements?” may be less threatening since it reduces the forcefulness of the statement.
17. **Do not ask respondents to make unnecessary calculations.**

Ideas for Constructing Questionnaires

To maximize the effectiveness of your survey effort, you also need to design the questionnaire in an effective manner. Simply writing down a set of questions and sending that out as a questionnaire may result in invalid or misleading data. Questionnaire design is an integral part of any survey effort.

In designing your questionnaire you should have two primary objectives: (1) reduce the amount of nonresponse and (2) reduce the amount of measurement error (Dillman (1999)). Nonresponse is when a potential respondent either makes no attempt to respond to any of the questions or leaves some questions unanswered. A well-designed questionnaire will encourage potential respondents to begin answering the questions and will ensure that no questions are missed. Measurement error is when respondents provide incorrect answers. It can occur when a respondent is confused about how to answer a question and can be caused by a poorly designed questionnaire.

Dillman's Thoughts on Question Ordering

Question ordering for surveys is a complex and often-discussed subject among survey designers. Dillman (1999) has provided the following six principles that you can use in ordering questions for your survey. One thing to consider when applying these principles is that they may contradict one another. That is, one principle may imply that a particular question be placed near the end, while another principle suggests that the same question be placed near the beginning of the questionnaire. In designing your questionnaire, you will need to weigh these competing influences.

Begin with questions that are the most salient to the respondent and work towards those that are the least salient. The respondent should have some association with the subject you are interested in, or that person would not be in the target population. Additionally, the cover letter that accompanies your survey should make a case that the subject matter is important. This means that respondents will expect to see and be more motivated by questionnaires that begin with questions that interest them. One implication of this is that you should not begin a questionnaire by asking a set of demographic questions.

Group similar sets of questions together. Respondents find it easier to answer similar set of questions when they are grouped together rather than scattered throughout the survey.

Place potentially objectionable questions near the end of the survey. An objectionable question may cause a respondent to discontinue the survey. This is less likely when such a question is at the end of the survey.

Order questions in a manner that will be logical to the survey respondent. You should consider how the respondent sees the subject matter and order the questions accordingly.

Group questions together that have similar components. For example, if you are asking a series of questions that require the respondent to rate their agreement or disagreement on a scale of one to five, then grouping these questions together may make sense. The reason is that the respondent can remain in the mindset appropriate for answering those types of questions.

Choose the first question carefully. The first question on a survey is the most important question in determining whether the survey will be returned or thrown away (Dillman, 1990).¹⁴ Dillman suggests the following in choosing an effective first question:

- (1) The question should apply to everyone that is responding to the questionnaire.
- (2) The question should be simple, requiring only a few seconds to read and answer.
- (3) The question should be something the respondent feels the need to answer.

Dillman's Thoughts on Constructing the Questionnaire

Once you have developed a set of questions and decided how to order them, you need to compile them into a questionnaire that will encourage the respondents to return the questionnaire. First, you need to create a questionnaire that creates the same stimulus for each respondent. Second, you need to construct the questionnaire in a way that focuses the respondent on answering the questions rather than on deciphering the questionnaire. This is often an overlooked aspect of survey design, but a poorly designed questionnaire can lead to high levels of both nonresponse and measurement error. Dillman has developed a set of 34 principles for constructing questionnaires. This guide does not repeat each of these, but rather includes a subset of 18 of those principles that EPA feels are most useful in conducting performance measurement surveys. These principles deal with both the visual design of the questionnaire and the placement of various pieces of information.

One common idea about improving response rates is that if the questionnaire is shorter, your response rate should be higher. This is true to some degree, but it will depend on how the questionnaire was shortened. Certainly, removing irrelevant and longer questions effectively reduces questionnaire length. On the other hand, reducing the font size or reducing spacing between questions may be counterproductive. These methods of reducing length may frustrate or confuse the respondent. Focus on designing the questionnaire effectively and logically, rather than on shortening the length through visual manipulation. Many of Dillman's principles will actually increase the length of a questionnaire by asking you to organize a question so that more vertical space is required for questions. His reasons for this, however, are based on research that has demonstrated that visual factors are more important than numbers of pages for influencing response. If you need to reduce the length of your questionnaire, focus

¹⁴ This is not to say that the first question is the most important *factor* in determining nonresponse rates, but is simply the most important of all the *questions*.

first on removing unnecessary questions and then on the types of answers you are requesting (e.g., open-ended questions usually require more space than closed-ended ones).

Table 4-1 presents the 18 principles that EPA feels are the most important in developing performance measurement surveys. They may seem somewhat repetitive, but the purpose is to provide a set of principles that cover all general situations, while also dealing with some specific ideas. As with some of the other principles above, satisfying all of these at once may not be possible since there are some contradictions.

Table 4-1. Subset of Dillman’s Principles of Questionnaire Construction

No.	
1	Write each question so that respondents do not need to re-read the question to know how to respond.
2	Place instructions exactly where the information is needed and not at the beginning of the questionnaire.
3	Use “item-in-a-series” lists to organize items with the same response categories, but be careful. An example of an “item-in-a-series” list would be asking respondents to rate, on a scale of one to five, a set of statements and then listing those statements in one column on the left with columns of ones through fives on the right. This is an effective manner of organizing questionnaires, but be aware that respondents tend to view these types of lists as comparative in nature and will respond to the set of items rather than to each question individually. That is, the answer to one question in the list will be made relative to all other questions in the list.
4	Ask one question at a time. For example, do not ask respondents to rate a presentation on a scale of one to five and then provide thoughts on how to improve the presentation.
5	Use similarity and diversity to identify groupings of visual elements. Questions should all look the same, similar response categories should all look the same, and instructions should all look the same. Additionally, you should make each group look different from one another (e.g., questions: bold; response categories: regular font; and instructions: bold italics).
6	Maintain simplicity, regularity, and symmetry to simplify the response task.
7	Number questions consecutively and simply, from beginning to end. That is, if you divide the questionnaire into sections, do not restart the numbering or use numbers such as “5.1.”
8	Leave more blank space between questions than between the subcomponents of questions.
9	Use dark print for questions and light print for answers.
10	Use question-specific instructions inside the question number and not as free-standing entities.
11	Separate optional or occasionally needed instructions from the question by a font change.
12	Do not place instructions in a separate book or in a separate section.

No.	Principle
13	<p>List answer categories vertically instead of horizontally, where possible. For example:</p> <p>How clear was the presentation of the material at the training seminar?</p> <p style="padding-left: 40px;">G Very clear G Somewhat clear G Somewhat unclear G Very unclear</p>
14	<p>Place answer spaces consistently to either the right or the left of answer category labels. For example, use either “Q Agree” or “Agree Q” and use the same order throughout.</p>
15	<p>Use numbers and simple answer boxes for recording answers. For questions that require a numeric answer, provide a blank line or empty box and provide a label for that answer (e.g., “___ pounds per month”). For closed-ended questions that provide pre-set answers, use a box such as this: Q.</p>
16	<p>Avoid double- or triple-banking of answer categories. For example, if you are providing 15 possible choices for a specific questions, <i>do not</i> arrange those in three columns of five. Either arrange them in one vertical column or reduce the number of choices.</p>
17	<p>Maintain the direction of scales throughout the questionnaire. That is, do not switch from displaying five-point scales as “5 4 3 2 1” to displaying them as “1 2 3 4 5.” Additionally, do not switch from having 5 as the “most positive” answer to having 1 as the “most positive” answer.</p>
18	<p>Use shorter lines to prevent some potentially key words from being missed. Visually, shorter lines tend to be read more completely than longer lines.</p>
<p>Dillman, Don, 1999. <i>Mail and Internet Surveys: The Tailored Design Method</i>, John Wiley and Sons, Inc: New York, Chapter 3.</p>	

B. Survey Implementation-The Tailored Design Method

This section discusses the Tailored Design Method (TDM) for implementing surveys. Dillman developed the TDM, which he describes in detail in his book *Mail and Internet Surveys: The Tailored Design Method* (John Wiley and Sons, Inc., 1999). The method is based on the principles of sociology and takes into account how people react and respond to questionnaires. This section first briefly discusses the five elements of the TDM and then discusses one of those elements, the stages of contact with potential respondents, in more detail. Appendix E contains sample letters of each stage of Dillman’s contact sequence that were used for the OC Metal Finishing program discussed in Section VI.

The Elements of the Tailored Design Method

Dillman's TDM comprises five distinct elements. Taken as a whole, these elements have been shown to achieve good response rates for mailed and other self-administered surveys. The five elements are:

- Element 1—Develop respondent-friendly questionnaires;
- Element 2—Implement a sequence of five contacts with the respondents;
- Element 3—Provide return envelopes with first-class stamps;
- Element 4—Develop personalized correspondence; and
- Element 5—Provide token prepaid financial incentives.

The first element (respondent-friendly questionnaires) is discussed in great detail in Section A. Element 2 (the contact sequence) is discussed in a separate section below. This section briefly discusses each of the other three elements.

Provide return envelopes with first-class stamps

Dillman suggests that the return envelopes that you provide for the respondents should contain first-class stamps. This personalizes the response process since stamps are less formal than metered envelopes. Research has shown that stamped envelopes are more effective than metered envelopes. Some respondents may very well take the stamps and not return the survey. Dillman, however, notes that respondents tend to be motivated by trust. In placing stamps on the return envelope you have said to the respondent, "we trust you will return this, but if not, here are some stamps."

Develop personalized correspondence

Surveys that look and feel like mass-mailings will tend to be treated like mass-mailings (i.e., many potential respondents will discard them). Dillman suggests that an effective method of avoiding this is to personalize each contact with the potential respondents. This includes signing each letter with an ink that differs in color from the letter's font color (e.g., a blue pen), using names in the salutation (e.g., Dear Ms. Smith), and writing letters using a familiar but businesslike tone. However, avoid over-personalizing correspondence since this can sound insincere.¹⁵

¹⁵ For example, "This new program will affect your shop, Smith's Auto Repair, and may result in a change in the way that you conduct exhaust repairs." may be too personalized by using the name and the type of business.

Provide token prepaid financial rewards

Dillman suggests that the use of token prepaid financial rewards can significantly increase response rates. He cites several studies that show that, by providing prepaid financial rewards, you can increase response rates by invoking a sense of trust in the respondent. However, in many cases financial incentives for returned surveys are prohibited by law for government-sponsored surveys, so you may need to think of alternative incentives (i.e., offering a palm pilot as was done for the Compliance Assistance Centers program).

Survey Implementation: Dillman's Five-Stage Contact Sequence for Mailed Questionnaires

Dillman (1999) proposes using a five-stage contact sequence to implement mailed surveys.¹⁶ He has found that this five-step process increases response rates by 20 to 40 percentage points over a single point of contact (e.g., a questionnaire with a cover letter only). Dillman advocates this staged process for two main reasons. First, the multiple stages act as reminders to those that would like to respond, but who have forgotten to do so. Second, the multiple stages offer the researcher several opportunities to convince a nonresponder to reply. The idea behind the staged process is to provide a new stimulus with each successive contact with the potential respondents. Using a new stimulus instead of simply resending previous materials is more effective, since people tend to discount repeated stimuli. The five stages of contact, along with their recommended timing, are:¹⁷

- 1) First contact—Prenotice letter;
- 2) Second contact—The questionnaire (4-7 days after the prenotice letter);
- 3) Third contact—Reminder postcard (one week after the original questionnaire);
- 4) Fourth contact—The first replacement questionnaire (two weeks after the reminder postcard); and
- 5) Fifth contact—Sending a final questionnaire through a different delivery mechanism (four weeks after the replacement questionnaire).

¹⁶ Dillman proposes financial incentives. However, since these are unlikely for government surveys, such incentives are not discussed in this section.

¹⁷ See Appendix E for sample letters for each stage.

Below is a detailed discussion of each of the five steps.

You should develop the materials for the first three stages at the same time. This will reduce logistical problems in administering the survey and allow you to develop a more systematic sequence. You can then wait until after you have mailed the postcards to develop the materials for the fourth and fifth stages. You may receive some information from respondents in the first three stages that will assist you in customizing these final two contacts.^{18, 19}

If you want to maximize the likelihood that your surveys are returned, you need to consider the schedule of your potential respondents. First, try to determine when potential respondents are likely to be busy and avoid those times. For example, mailing questionnaires to farmers at harvest time will likely result in significant nonresponse. Second, try to determine when your survey will be most salient to the respondents. For example, it is better to mail a survey on auto repair shop procedures for fixing air conditioners in the summer than in the winter. Finally, avoid mailing the survey so that it arrives between Thanksgiving and Christmas since this is always a busy time for people.

First Contact—Prenotice Letter

The prenotice letter tells respondents that they will be receiving a survey and that you would appreciate their assistance by completing and returning it. The letter should be brief (one page) and should not raise any concerns for the respondent. For example, avoid long discussions of confidentiality. The letter should contain each of the following components:

- # The date;
- # The potential respondent's address (i.e., an inside address);
- # A first paragraph that states the potential respondent will be receiving a survey and who is sponsoring the survey;

¹⁸ For example, you may find that your sampling frame included a number of out-of-scope facilities. This may be something to address in subsequent contacts.

¹⁹ If you need OMB approval for your questionnaires, it may be more efficient to include all five letters in the initial approval request.

- # A paragraph that explains the subject of the survey;
- # A paragraph that explains why the survey is useful and what the results will be used for;
- # A thank you phrased in sincere terms; and
- # A real signature.

Finally, print the letter on the letterhead of the sponsoring Agency and mailed it first class. Time the letter so that it arrives only a few days ahead of your questionnaire.

Second Contact—The Questionnaire

Your second contact will include three items: a cover letter, the questionnaire, and a stamped return envelope. Earlier sections discussed the questionnaire and the use of stamped return envelopes, so this section discusses the cover letter for this mailing and also provides some tips on this stage of contact.

The cover letter for this mailing will differ from the prenotice letter by providing more detailed information. Nevertheless, keep the cover letter to one page since longer letters will only raise concerns among potential respondents. Your cover letter should include the following components:

- # The date.
- # The potential respondent's address (i.e., an inside address).
- # A first paragraph that explains why the respondent is receiving the questionnaire. It should explain in general terms the subject matter of the survey. Begin the paragraph with an expression such as "I am writing to ask for your assistance in a study...".
- # A paragraph that explains how and why the respondent was selected to receive a questionnaire. This paragraph should explain who the target population is and how members were selected from that group (e.g., randomly selected).
- # A paragraph that explains the usefulness of the survey. Specifically, you should discuss how you will use the results. If you can relate the usefulness of the results to the respondent, do so. For example, for developing a compliance assistance program, include a statement such as "We will use the results of this study to develop a program that best fits your needs."²⁰

²⁰ The example statement is very general; strive to be more precise with your own statements.

- # A paragraph that discusses confidentiality. Keep this paragraph simple and to the point. A long discussion about confidentiality may raise concerns among some respondents. Some respondents may not be convinced about confidentiality of the responses, but you cannot direct your letter solely at those members of the population. In a subsequent paragraph, provide a phone number where additional questions can be asked.
- # A paragraph that indicates you are willing to answer any questions regarding the study, the questionnaire, or confidentiality. Provide a phone number and the hours when someone will be there to answer questions.
- # A sentence thanking the respondent for taking the time to complete the survey.
- # A real signature.

As with the prenotice letter, print this letter on the sponsoring Agency's letterhead.

Finally, below are some additional tips on the second contact:

- # *Use questionnaire identification numbers*—Each questionnaire should have unique identification (ID) number. Print the number in the upper right corner of the questionnaire cover page. This will assist you in tracking who has responded.
- # *Mail the questionnaire packet using first-class postage*—Although this is more expensive than other bulk means of delivery, Dillman's experience is that the benefits of using first class far outweigh the costs.
- # *Review postal procedures*—The U.S. Post Office changes postal procedures from time to time. Before sending out a large mailing, you should review the most current postal procedures and inform the local postal office of the size of your mailing. They may be able to better accommodate you and provide some advice to ensure a successful mail-out.
- # *Put some thought into assembling the package*—This can often be overlooked. All items in the package should come out at once when the respondent opens the packet. Try different ways of packaging the components and then ask someone who does not know what to expect inside the envelope to open and remove the items. Also avoid using paper clips to hold the materials together since they can jam postal sorting machines.

Third Contact—Reminder Postcard

This stage involves the final remainder that is sent to all members of the sample. One important thing to remember about this stage is that using a postcard is not simply to save resources but to provide a new stimulus for the potential respondents. The primary purpose of the postcard is to get responses from

those individuals who want to return the survey, but have yet to do so (e.g., have forgotten about or lost the survey). The postcard will have little impact on “hard-line” nonrespondents. The fourth and fifth contacts are directed at them.

Your postcard should contain five items:

- # The date;
- # A first paragraph that explains a survey was sent to them, when it was sent, and what the survey concerned;
- # A second paragraph that (a) thanks the respondent if they have already completed and returned the survey, (b) encourages the respondent to complete and return the survey if they have not done so already, and (c) a reiteration of how important each respondent is to the study;
- # A final paragraph that gives a phone number where they can request a replacement questionnaire; and
- # A signature.

The postcard should use the same name and address as the previous mailings. You do not need a salutation on the text side of the card.

Fourth Contact—The First Replacement Questionnaire

This mailing will be similar to the second contact, but make your cover letter more insistent and send this only to nonrespondents. Include new copies of the questionnaire and stamped return envelope. Simply resending the original questionnaire packet and stamping “second notice” on the materials will be ineffective. Those that did not respond to the initial request will probably not complete an exact copy. In this stage, your cover letter must convey to the nonrespondents the importance of their response.

Your cover letter should include the following information:

- # The date.
- # The potential respondent’s address (i.e., an inside address).

- # A first paragraph that states you sent them a questionnaire, how long ago it was sent, what the questionnaire concerned, and that you have not heard from them. Be direct and sound as if you are singling them out. Although this sounds confrontational, it is an effective means of obtaining responses. This first paragraph must get and hold their attention to read the remainder of the letter. Remember that the people receiving this letter have not yet responded and bold statements may be needed to get their responses. You will soften the confrontation nature of this communication as you go along.
- # A second paragraph that explains that others have responded. The idea here is make it known that it okay to respond since others have done so. That is, their response is not being singled out, just their participation.
- # A third paragraph that reiterates how important their response is to the study. Convey that their response will add valuable information to the study.
- # A fourth paragraph that gives them a way out of future mailings if they are not part of the target population. For example, ask them to return the survey and state briefly why they are not eligible.
- # A fifth paragraph that reiterates your commitment to maintaining confidentiality. Nonrespondents at this stage may have this concern, so assure them that their responses are confidential.
- # A sixth paragraph that conveys an understanding that some people may not want to complete the survey. Ask them to return the questionnaire with a brief note on the front explaining why they do not want to fill it out, or to simply return it blank. In return, offer to remove them from future mailings. The information you obtain here may help you understand the nature of your nonrespondents.²¹
- # A real signature.
- # A postscript that provides the phone number to call with questions.

As with the previous letters, print the letter on the letterhead of the sponsoring Agency.

Fifth Contact—Sending A Final Questionnaire Through A Special Delivery Mechanism

The final contact provides a new stimulus by using a special delivery mechanism (e.g., certified mail) or a different method (e.g., telephone call) to get the questionnaire into the hands of nonrespondents. This mailing will contain a new cover letter and new copies of the questionnaire and stamped return envelope.

²¹ You can also combine the fourth and sixth paragraph, especially if the remaining number of nonrespondents is small.

Once again, revise the cover letter to reflect the stage of contact. Specifically, your cover letter should include:

- # The date.
- # The potential respondent's address (i.e., an inside address).
- # A first paragraph that states you have been sending them materials and what those materials have concerned (i.e., a study and the subject matter).
- # A second paragraph that states the usefulness of the study.
- # A third paragraph that explains "time is running out" for them to respond.
- # A fourth paragraph that explains you have sent these materials via special mail service (state which one specifically) to ensure that they receive the material because all responses are important in producing reliable results.
- # A fifth paragraph that explains the confidentiality of the survey.
- # A sixth paragraph that thanks them for their participation. This should be more than one sentence.
- # A real signature.

Print the letter on the sponsoring Agency's letterhead.

Given the variety of delivery options, consider which one will work best for you. Among your choices are:

- # U.S. Postal Service certified mail;²²
- # U.S. Postal Service priority mail; and
- # Courier services (e.g., Federal Express, United Parcel Service).

If you have several items to mail, contact the service you will use and inquire about any procedures you should follow to ensure timely delivery.

Tracking Techniques

²² Use this option only if you can ensure that someone will be there to sign for the package.

Survey tracking is a valuable part of your data collection effort. You can track surveys both in regards to the population to which they apply and the time frame in which they are mailed and returned. Population tracking can be helpful when sending surveys that are to be returned anonymously. By using general codes printed on the survey, you can track the returned information by strata, such as state or type of facility, without requiring the respondent to identify themselves through name or address. Mailout and delivery tracking is used to track the progress of survey mailout and response activities. This information can help determine the number of respondents from each step of the TDM contact sequence, as well as the associated return rates. You can develop simple tracking systems using commercially available spreadsheet or database programs. Keeping tracking information in an electronic media such as a spreadsheet or database helps simplify analysis.

Data Entry and Handling

Once surveys are returned and logged into the tracking system, enter the data into some type of electronic media. If the data are to be analyzed in multiple ways, the best option is to use a database application where querying the data is relatively simple. If you will conduct statistical analysis on the data, spreadsheets are effective. In most cases, data can easily be converted from a spreadsheet to a database and back again if necessary.

Follow-up and Clarification

Some survey recipients may not completely fill out their forms or may provide contradicting information. Other respondents may not write clearly. In these cases, if the survey was not anonymous, you may need to follow up with the respondent to clarify the answers. Follow-up is most successful when you can contact the respondent via telephone. In these calls, be as courteous and brief as possible to keep the additional burden to the respondent to a minimum.

C. Where Do We Go From Here?

This section provides some basic information on analyzing the data that you have collected. Following are definitions that will help you understand some basic analytical techniques. We then discuss

some basic analytical techniques, separating them into descriptive techniques and measures of association and comparison. Appendix F expands on these concepts and presents advanced analysis techniques.

The following subsections provide details on how to implement a number of statistical techniques as well as what those techniques will tell you. Below are some definitions that will help you understand the discussions that follow.

Variable—A variable is some measure of the characteristics of your sample or population. Examples of variables are the pounds of chemicals used by each facility, the compliance status of a facility, and the answers to survey questions. Variables are what statistics analyze.

Limited Variables—A limited variable is one that can only take on values within a certain range. For example, when coding a **yes/no** question, you usually code “yes” as 1 and “no” as zero, thus limiting the variable to two values. A question that asks respondents to rate satisfaction on a scale of one to five is also a limited variable. The opposite of a limited variable is an “unlimited” one, or a variable that has no constraint on the range of values. The pounds of chemicals used by a company could be considered an “unlimited” variable.²³ In general, the term “unlimited” is not used to describe variable; rather, it implies that the variable is not considered limited.

Continuous Versus Discrete Variables—These are ways of describing the potential values that variables can take on within the feasible range of values, where the feasible range of values is defined by the variable’s limits. A discrete variable has specific values, usually integers, that it can take on. Examples of a discrete variable are yes/no questions and questions that ask respondents to respond on a five-point scale, since the numerical answer is limited to a few values. Although many discrete variables are also limited (see above), some are not. For example, a question that asks how many pieces of a certain type of equipment that a facility has in use is discrete, but not limited: (i.e., the variable can only take on integer values, but there may be no limit on the number). A continuous variable can, theoretically, take on any numerical value. In general, if the variable can take on decimal values (i.e., non integer values), then it is continuous.

Scaled Variable—A scaled variable is a discrete variable that has been given a set of values that correspond to qualitative criteria. For example, a question that asks respondents to rate their satisfaction with something on a “scale” of one to five where each value has a defined level of satisfaction is a scaled variable.

Statistical analysis procedures are generally separated into either descriptive techniques or inferential techniques. This section discusses some basic descriptive techniques you can use to describe your data. The purpose of descriptive statistics is to provide a feel for the nature of the data.

²³ Granted, firms cannot use less than zero pounds of a chemical, which effectively “limits” this variable. However, limits such as this are not generally considered limits unless a large number of facilities use zero pounds.

Frequency Distributions

A frequency distribution summarizes of how the data are distributed across various numeric values for certain characteristics or factors in the data. The distribution can tell you where the data are concentrated and what values are uncommon in the data. Below is an example simple distribution for the size of facilities responding to a surveys:

Table 4-2. Example Frequency Distribution

Number of Employees	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1 - 100	15	12.8%	15	12.8%
101 - 200	56	47.9%	71	60.7%
201- 300	34	29.1%	105	89.7%
301 - 400	12	10.3%	117	100.0%
TOTAL	117	100.0%	-	-

The table shows that the largest number of facilities employed between 101 and 200 people and that 60.7 percent employed 200 or less.

This example contains the four standard elements of a frequency distribution:

- # *Frequency*—The number of entities for each category;²⁴
- # *Relative frequency*—The percentage of the total represented by the frequency for that category;
- # *Cumulative frequency*—The sum total of the frequency for each category and the frequencies for categories that have been listed above it; and
- # *Cumulative Relative Frequency*—The sum total of the relative frequency for each category and the relative frequencies for categories that have been listed above it.

²⁴ In this case, it is a range of employees.

Although you do not have to calculate each of these for your data, you should always report the frequency. The “cumulative” measures have little meaning when the categories are nonnumeric or unordered (e.g., counties in a state). For the most part, your data and informational needs will dictate what you calculate.

You can also use charts rather than tables to represent frequency distributions. The bar chart to the right in Figure 4-1 represents the same employment distribution as is shown in



Figure 4-1. Example Chart of Frequency Distribution

Table 4-2. Furthermore, there are numerous charts that one can use to convey data. For example, the employment distribution data can also be represented by a pie chart as shown in Figure 4-2.

Both the two charts and the table provide the same information (i.e., the relative size of each employment class), but do so in slightly different ways. For the most part, which one to use falls to personal preference. You should remember, however, that the table representation provides more information and may be more suitable for a report. On the other hand, charts tend to work better in presentations.

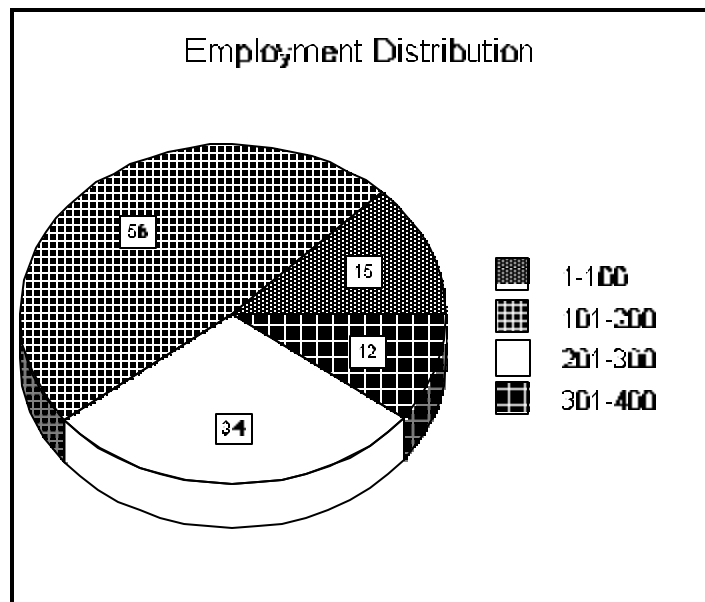


Figure 4-2. Example Pie Chart for Frequency Distribution

For numeric data, there is no standard method of defining the ranges that

form the categories (e.g., the 1 - 100, 101 - 200, 201 - 300, and 301 - 400 employee categories). In some cases, the categories can be based on well-defined groupings (e.g., counties within a state). In other cases, the categories are more difficult to define. The example uses equal-sized categories (100 employees per category). Real-world situations are rarely this convenient. For example, in the case study in Section VI of the guide, we divided the metal finishers into two unequal employment size classes: less than or equal to 10 employees and greater than 10 employees. In short, you may need formulate your distribution a few times before you determine a set of classes that provides the information that you want.

Charts are also useful for presenting information on a number of questions simultaneously while also providing comparative information for those questions. Consider the following set of questions where the usefulness of a set of resources are evaluated by respondents.

Q2. Please indicate the usefulness of these activities. Rate each using the following scale: 4–“very useful”; 3–“somewhat useful”, 2–“somewhat not useful”, 1–“not useful at all”.	
EPA Presentation for colleges and universities	1 2 3 4
Other EPA presentations for broader audiences	1 2 3 4
EPA audit protocols	1 2 3 4
EPA lab guide	1 2 3 4
Frequently Asked Questions Featured on EPA web site	1 2 3 4
Other EPA website information	1 2 3 4
EPA responses to phone calls/hotlines	1 2 3 4

Questions of this sort provide information on the usefulness of individual resources, but also provide information on the relative usefulness of the various resources. Figure 4-3 provides an efficient means of providing both levels of information (individual and relative usefulness) in one chart.

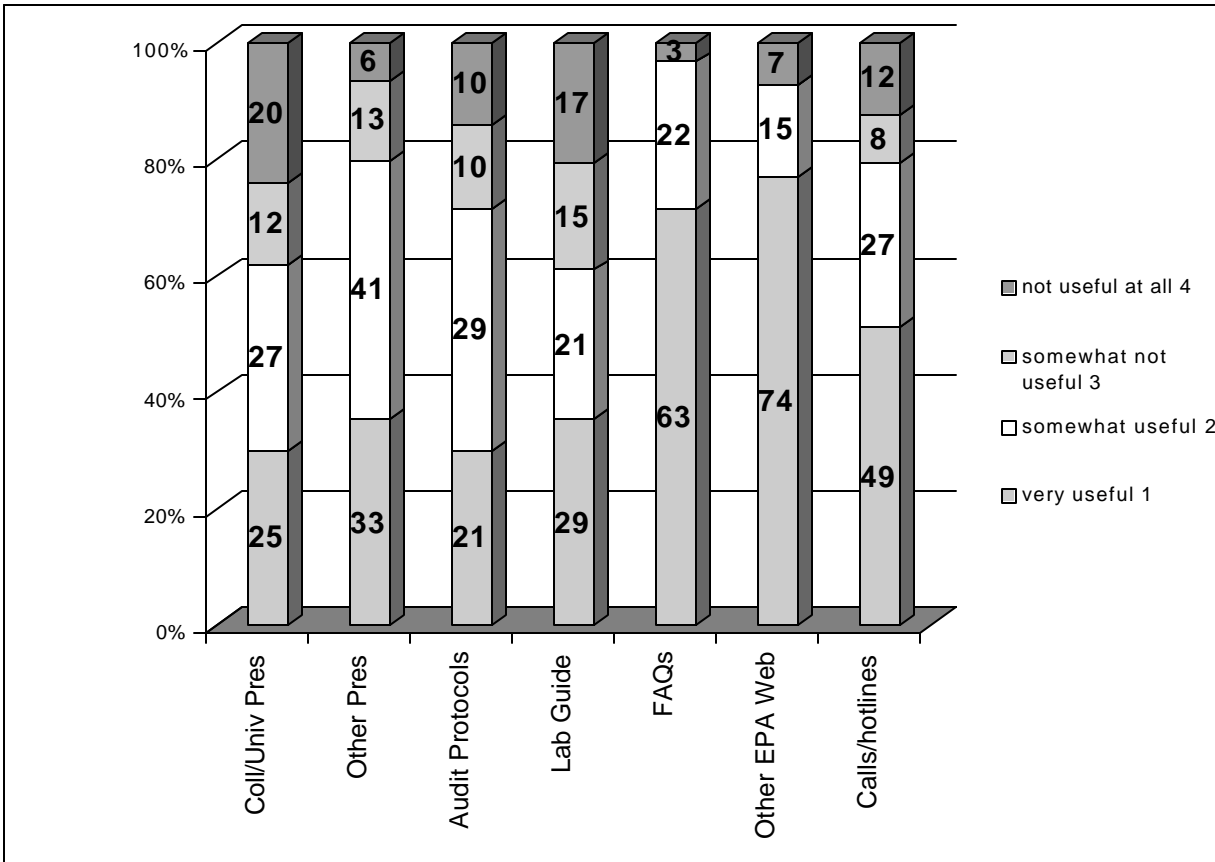


Figure 4-3. Example of Answer Distribution Chart

Measures of Central Tendency—Mean, Median, and Mode

A measure of central tendency attempts to provide a single-value description of where the values in the data are centered. In other words, if we were to pull a number from the distribution, what value is it likely to be closest to. The most common are:

- # Mean—The average value of the data;
- # Median—The value that has half the values above it and the other half below it; and
- # Mode—The most common value in the data, or the value that appears most frequently.

Each has strengths and weaknesses depending on the nature of the data.

The mean is almost always calculated and is generally a good measure of central tendency, especially when the data are not highly skewed toward either low or high values in the data. The mean, however, can be easily influenced by one or two outlier observations (i.e., observations that are either very high or very low compared the rest of the distribution), thus skewing the estimated mean toward the outliers. Means do not work well when the answer to a survey question is not numeric (i.e., a qualitative answer).

The median is also almost always calculated for a set of data.²⁵ Additionally, the median generally overcomes the shortcomings of the mean (i.e., it works well when data are highly skewed since it is not influenced by outliers). The median, however, does not easily lend itself to further statistical analysis, thus limiting its use beyond descriptive purposes.

The mode is the value that appears most often in the data. The mode is not used as often as the mean and the median and is relevant when the data are known, discrete values. For example, a survey question that allows respondents five possible responses would be well described by the mode. Variables that are more continuous, such as pounds of chemicals used, are not well described by the mode since the data could, theoretically, take on different values for each respondent.²⁶

These three measures are not substitutes for one another, but are three different measures that can describe the data differently. Therefore, you could calculate all of them or only one or two.

Measure of Variability

A measure of variability tells you the extent to which the data values differ from one another. In some sense, the frequency distribution provides this type of information. The measures described here, however, summarize variability using a single value rather than a set of frequencies for different data

²⁵ Most spreadsheet programs contain a median function.

²⁶ One exception to this, however, is when a number of respondents report the same value. For example, in a sample of 100 entities, 20 may state that they use zero pounds of a certain chemical, with the remaining 80 report 80 different values. In this case, the modal value of “zero” is a useful piece of information.

ranges. For simple analyses, there are two measures of variability that you should be concerned with: the range and the variance.²⁷

The range is simply the difference between the largest and smallest value in the data. The range is useful when you want to demonstrate that respondents reported either a wide range or a narrow range of values for certain variables.

The variance and the standard deviation²⁸ measure the extent to which the values in the data differ from the mean value. You can calculate the variance of a sample by summing the squared differences between each observation and the mean value and then dividing by the number of observations minus one. Most spreadsheet programs contain a variance function. Larger variances indicate the data are more disperse.

You can also use the standard deviation in situations when there is a normal distribution.²⁹ If you can assume that the population from which the data were drawn are normally distributed, then the following rules hold:

- # 68 percent of all observations are within one standard deviation of the mean;³⁰
- # 95 percent of all observations are within two standard deviations of the mean; and
- # 99 percent of all observations are within three standard deviations of the mean.

These rules allow you to infer the overall distribution of the population.

Two Other Useful Descriptive Measures

²⁷ In actuality, there are three measures. The third, which the standard deviation, is derived from the variance.

²⁸ The standard deviation is simply the square root of the variance.

²⁹ A normal distribution is also referred to as a *bell curve*. For normal distributions, the mean and median are equal and the distribution above the mean mirrors the distribution below the mean.

³⁰ Because we are assuming a normal distribution, the percent ages (68, 95 and 99) will be evenly divided on either side (above and below) of the mean.

Finally, there are two other useful descriptive measures: minimum and maximum values. These values provide a useful estimate of the smallest and largest values for the underlying population.

D. Lessons Learned from Measurement Projects

One of the primary goals of this revised guide is to share lessons learned from the Regional pilot projects and other sources. The lessons and experiences learned since the first version of the guide was published in 1999 are reflected and incorporated throughout this document. This section lists some of the notable lessons learned in the areas of general program evaluation, questionnaire design, and questionnaire implementation, in order to give users a sense of the experiences of the Regions. Most of these lessons relate to key messages of this guide: plan carefully, know the goals of your compliance assistance and the purpose of your evaluation, use an appropriate collection tool, and know your target audience.

General Program Evaluation

- # **Keep measuring reach and outputs.** EPA emphasizes outcome measurement because measuring for results provides the best indication of how well the Agency is protecting human health and the environment. Outcome measurement, however, does not always provide a complete picture of the quality of a particular activity. Continue to measure reach to ensure that you are reaching your target audience. Also, continue to measure outputs (e.g., the number of compliance assistance guides distributed, the number of workshops conducted, the number of site visits conducted), because outputs show how well the program is fulfilling its responsibility to reach out to its target audience. In addition, where outcome data are unavailable or unclear, output data may be the best or only source of performance information.

Questionnaire Design

- # **Ask respondents to identify their position.** It's not uncommon for individuals who are not affiliated with a target facility, such as secondary compliance assistance providers, to receive compliance assistance tools, attend compliance assistance presentations, or use Web sites and helplines. Some of the pilot projects showed that, as respondents, these individuals can skew results to make the compliance assistance appear less effective. They automatically answer some questions in the negative because they are not in a position to make a behavioral change or environmental improvement. If you cannot eliminate these individuals from the respondent pool, consider asking them about the nature of their position, so that you can sort your results appropriately.

- # **Take language issues into account.** Some of the pilot projects had a target audience that included many non-English or English as a second language speakers. One pilot's measurement efforts showed that it is important to have high quality translations of compliance assistance and survey materials to avoid confusion and maximize effectiveness. Translation services do require additional resources, however, and that needs to be considered in your project planning as well.
- # **Match pre-/post-test questions to the compliance assistance material.** If you know you want to measure understanding through pre-/post-tests, develop your compliance assistance material in advance, to allow adequate time to put together matching test questions. If your compliance assistance presentation, seminar, or workshop does not follow a predetermined structure—for example, you are conducting a question and answer workshop—or is formulated soon before it is delivered, you may not be able to develop test questions that accurately assess the effectiveness of your efforts.
- # **Do not use comment cards.** EPA and states found that these cards were potentially biased and received very low response rates (usually around 2 percent). Such a low response rate does not make the evaluation effort worthwhile and you should use another data collection tool for your measurement projects.

Questionnaire Implementation

- # **Select an appropriate time-frame for conducting the evaluation.** Some of the pilot projects were affected by the high turnover rates in some industries. For one project, waiting just six months between delivery and measurement of a compliance assistance project resulted in a much lower response rate from employees leaving facilities. If you are measuring changes in understanding, you might want to follow up sooner because of the possibility of respondent turnover. At the same time, if you are measuring behavioral changes or environmental and human health improvements, you will want to conduct the evaluation after enough time has passed for changes to be implemented. If possible, try to periodically follow up with recipients of compliance assistance.
- # **The Tailored Design Method works.** Regional pilot results showed that the TDM can produce significantly higher response rates for mailed surveys. Past pilots have also shown that surveys that are not well-planned can have very disappointing results (past OC response rates have been approximately 17%). The TDM suggestions in the guide were selected because of their potential as resource-effective ways to get the most out of your surveys.
- # **Direct observation works for measuring behavioral change.** Several pilot projects reported high satisfaction with the use of direct observation during onsite revisits. Direct observation helped in making accurate assessments, measuring compliance changes, obtaining detailed feedback from respondents, and observing innovative practices. Being on site also helped compliance assistance providers learn about the language and culture of their target audience and identify areas where future assistance might be provided.

- # **Get the most out of your pre-/post-tests.** Several pilot projects showed that some attendees will skip post-tests at the end of a workshop because they are tired or need to depart for travel reasons. If possible, try to schedule pre-/post-tests immediately before or after compliance assistance presentations when attendees are already assembled. You should also let attendees know how long the tests are expected to take, so they do not need to worry about committing too much of their time.

SECTION V: AN INTRODUCTION TO STATISTICAL SAMPLING

This section provides basic information on how to draw a statistically valid random sample to measure the performance of compliance assistance activities. After reviewing this section, you should be familiar with a number of statistical concepts, be able to estimate the necessary sample size needed to ensure statistical validity, and know how to draw the sample from your population. Please note that you cannot use the generic ICR for a statistical evaluation; you will have to develop a separate ICR for any statistical evaluations. (See Section III E for more information about obtaining OMB approval and using or developing an ICR).

The information provided here is simplified and should not be considered a replacement for the services of qualified survey statistician. This section is designed to assist you in developing a sampling plan for a project that has few complications or special circumstances. Appendix F contains information on advanced data analysis. For more complicated projects, you should consider using a qualified survey statistician. Furthermore, you should consider having a statistician develop or review the sampling plan for even simple projects where you will use the results of the survey to support important policy decisions. Use the information in this chapter when:

- # You know the population from which the sample will be drawn; Appendix G contains information to use when sampling from unknown populations.
- # You need to measure specific trends or aspects among the *whole population* (e.g., use of compliance assistance tools among auto repair shops) or you need to *divide the population into subsets* and look at trends within or among those subsets (e.g., use of compliance assistance by *independent* auto repair shops and by *shops that are part of dealerships*); and
- # Your main interest can be reduced to a “yes/no” binary-type question (e.g., in or out of compliance, found the compliance assistance project useful or not).

This section identifies potential pit-falls, complications, and considerations that could warrant the use of a qualified statistician.

This section presents definitions of some basic statistical terms, followed by a brief definition of statistical validity. Next are aspects of a sampling plan that you will need to consider: the sampling frame, the sample size estimate, and the sampling method.

A. Key Terms

Below are definitions of key statistical terms that you should know if you are going to conduct a statistical sampling project. These terms will be used throughout this section.

- # **Population**—The complete group of entities that you wish to analyze. In terms of compliance assistance programs, the scope of the program should define the population for a sampling plan.
- # **Sample**—The subset of a population that participates in the survey. Taking a sample, instead of measuring the whole population, reduces the cost of analyzing the whole population. A measurement based on a sample is less accurate than if you measured the whole population, but by adhering to statistical sampling principles you can obtain fairly accurate estimates and, to some extent, control the resulting degree of inaccuracy.
- # **Precision**—The amount by which we are willing to be wrong in our measurement. A good example of precision comes from political polling with statements such as “45 percent of people polled will vote for candidate X, with an error of 4 percentage points.” This means that 45 percent (plus or minus 4 percentage points) will vote for candidate x (a range from 41 to 49 percent). The “4 percentage points” is the precision of the estimated 45 percent. When you draw a random sample, you *choose* a level of precision for your estimate, which is the amount by which you are willing to be wrong and still be comfortable with your estimate.
- # **Confidence Level**— The probability that you are within your specified level of precision. The actual percentages are based on statistical equations that won’t be discussed here. Confidence levels are generally set at 90 or 95 percent. Although 95 percent confidence levels increase the precision of your estimates relative to a 90 percent level, where resources constrain the sample size that can be collected, EPA is comfortable with use of a 90 percent level.
- # **Unit of Response**— The entity about which you are collecting information. For the most part, a unit of response for compliance assistance performance measurement projects is the facilities where the assistance is directed.
- # **Bias**—A sample is biased if the entities in the sample are not representative of the population as a whole. For example, if a population of facilities that you are surveying is evenly divided between small and large facilities and you collect data only from large ones out of convenience, then the sample is biased. In collecting a sample, the idea is to

evaluate the population as a whole. To do this with some accuracy and confidence, you must be assured that the sample is not biased in any way.

B. Statistical Validity

In drawing random samples to measure performance or compliance, you must follow standard statistical techniques to ensure that the resulting measures and/or comparisons are statistically valid. In statistics, a valid study measures what you intend it to measure. Specifically, you can obtain a statistically valid sample by using sampling procedures that result in a known precision and confidence for the resulting measure and that obtain an unbiased estimate of the measure. Section D discusses how precision and confidence requirements impact your choosing a sample size. In general, by adhering to statistical sampling procedures, while also understanding the limits of these procedures, you should obtain statistically valid measures.

C. Sampling Frame—The List of Potential Respondents

The sampling frame is the list of entities from which you draw your random sample. In its simplest form, a sampling frame is a complete list of entities that you are interested in. Each entity on the list will be “in scope” and no entities that might interest you are excluded. Developing an accurate and complete sampling frame is one of the most important, and often overlooked, aspects of sampling.

The first step in developing a sampling frame is to define the scope of your effort. Usually, this should match entities at which you have targeted the compliance assistance effort. You may be interested in only the subset of the facilities that received compliance assistance (e.g., those that attended seminars, to gauge the impact of the seminars). What is most important here is to develop a concise definition of the facilities that are “in-scope.” For each possible entity it should be clear whether or not it fits into the scope.

In developing your sampling frame, you should understand the distinction between the *target population* and the *sampling population*. The target population is the group of entities in which you are interested and is defined by the scope of the effort discussed above. The sampling population is the group of entities from which a sample is pulled. If a sampling frame is perfect, then the target population

matches the sampling population exactly. This is rarely the case, however. Developing a good sampling frame involves minimizing the differences between the sampling population (defined by the frame) and the target population (defined by the scope).

Differences between the target and sampling populations can result in serious biases that may impact your results. Consider a survey that looks to measure compliance among dry cleaners (target population) by drawing a random sample of dry cleaners that are participating in a voluntary program (the sampling population). Unless all dry cleaners are part of the voluntary program, this survey will not be able to make unbiased inferences about all dry cleaners. That is, we might expect that those who have volunteered to be in the program will differ from those not in the program. If the sampling frame and the target population do not match, then you will not be able to conduct unbiased analyses of your target population. In actuality, in any sampling effort, you can only make unbiased inferences about your sampling population since that is where you drew the sample from. Thus, it is important to have the target population match the sampling population as best as possible.

EPA suggests that once you have developed your sampling frame, take some time to determine how well your list matches your target population. Here are some guidelines to develop and review your sampling frame:

- # **Coverage**—How well does the sampling frame cover the target population? Does the sampling frame include some entities that are not in the target population? If so, remove them. Does the sampling frame omit any entities in the target population (e.g., a trade association list may not cover all entities)? If so, can you add those entities?
- # **Duplicates**—Are there duplicates within the sampling frame? That is, does each entity appear only once in the list? If entities appear more than once, remove the duplicates.
- # **Current**—How current is the list? Is the list several years old or was the list developed recently? Obviously, the more current, the better.
- # **Information Content**— Does the sampling frame contain enough information for each entity for the survey method that you will be using (e.g., telephone number, mailing address, email address)? A frame that perfectly matches the target population, but is missing addresses for several entities, will not support a mailed survey. Using such a frame for a mailed survey and simply dropping those without addresses amounts to changing, and potentially biasing, the sampling population to “those with addresses in the sampling frame.”

- # **Stratification Criteria**— If you are going to divide the sample into multiple groupings (e.g., large, mid-sized, and small entities), you will need to ensure that your frame contains the information to make those divisions. For example, you may want to divide a set of auto repair shops between independent shops and dealer-based shops. You will need to know into which category each shop falls. You should avoid assuming that entities can be placed into certain groups without definitive information to back up that assumption. This be a serious issue where the frame is derived from more than one source and the different sources contain varying levels of information.³¹

The sources for your sampling frame should be reliable and should provide all of the information that you will need to determine which entities are “in scope” (e.g., entities that use a specific process) and to administer your survey. For compliance assistance performance measurement projects, sampling frames can come from sources such as:

- # Regulatory information;
- # Trade associations; and
- # Directories.

You should document the source for each entity in your final sampling frame. This information may assist you in identifying biases after the survey has been completed.

You cannot always develop a sampling frame that closely matches your target population at the beginning of the project. For example, suppose you are interested in determining compliance and the need for compliance assistance among metal finishing facilities performing chrome electroplating. If there are no reliable publications that document which facilities perform this operation, then developing an accurate sampling frame at the start would be impossible. You can, however, use a sampling frame that includes all metal finishers and then incorporate a screening stage into the sampling plan. In the screening phase, you would contact entities to determine if they are “in scope” before they are asked to be a respondent. This type of situation, however, is best handled by an experienced survey statistician.

An even more complicated situation would be where there is no reliable list of potential respondents (i.e., the population is “unknown”). For these cases, we have provided some guidance on sampling from unknown population in Appendix G. There are statistical sampling methods that can be

³¹ A related issue for stratifying multiple-source frames is that two sources may contain the same data element (e.g., employment levels), but one source’s data may be more accurate or more reliable than the other sources being used.

used to develop sample frames (e.g., cluster sampling), but implementation of these methods should be left to an experienced survey statistician.

Finally, you should also consider if you may want to stratify (i.e., divide) the population. Stratification can be done for two reasons. First, you may want to make valid statements about some subgroups of the population (e.g., a statistically valid compliance rate for small facilities). This is discussed in more detail in Section D below. Second, you may simply want to ensure that some facilities of a certain type are included in the sample (e.g., ensuring that the sample includes some small and some large facilities when large facilities represent a small proportion of the population). Finally, you should also consider how you may want to stratify (i.e., divide) the population. Specifically, your sampling frame will need to contain the necessary information to stratify the sample. For example, if you want to stratify by size (e.g., small versus large), you will need information on the size of each facility in your sample. This can be a serious issue when you have used more than one source to develop your sampling frame. You need to make sure that each source contains the necessary information. Additionally, you also need to make sure that the information contained in each source is consistent. For example, a source that provides specific numbers of employment size (e.g., 5,115) is probably not consistent with a source that provides rounded numbers (e.g., 5,000)³². This has to be judged in context, however. In the above example, you should review the number of facilities that may be close to the cut-off to determine the likely “error” associated with inconsistent sources. A bigger issue would be where one source provides the necessary information and another does not. In this situation, stratification may not be possible.

D. Estimating A Sample Size

You can estimate the sample size for a simple survey by answering six questions and then applying one of two formulas. This section discusses those six basic questions and provides the two basic formulas along with information on when each is applied. The method for estimating a sample size presented here assumes that your survey is simple and straightforward. The discussion that follows points out situations in which you may want to turn to a qualified survey statistician. Furthermore, if your answers to these six basic questions do not fit neatly into the possibilities that this section provides, or if you are unsure about your answers, you may consider using a qualified survey statistician.

³² Rounding the specific source to match the rounded one may not solve the inconsistency problem. The problem may stem from using different original sources for the employment numbers.

Question 1: Are you measuring a proportion or an average value?

This question involves the nature of what you are measuring. To answer this question, think about the factor that you are most interested in. Is it a compliance rate? Is it the use of a specific chemical? Naturally, you may want to know several things about the entities, but what is the most important piece of information? Ideally, this will be tied directly to your compliance assistance program. Once you have determined that, can it be described as a proportion (e.g., the percentage of entities in compliance, the percentage of entities using a certain chemical) or an average value (e.g., the average amount of a chemical used)? In general, a proportion will be a yes/no question for the entities in the survey while an average will be a measure that can be one of many values or a range of values for each entity. For simple surveys, proportions are much easier to handle. For that reason, the remainder of this section will only deal with sample sizes that are based on measuring proportions. Estimating a sample size for average values involves some additional complications that may be best handled by a survey statistician.³³

Question 2: Are you concerned with measuring whether you are above (or below) a specific value for your factor, or are you concerned with simply measuring the value of the factor?

How to answer this question is best illustrated by example. Suppose you have some information (e.g., self-reported data) on the compliance rate in a sector and that information suggests compliance is at 90 percent. You may want to verify compliance is actually that “high” using a statistical sample. In this case, you are concerned with measuring whether or not the factor (compliance rate) is below a specific value. Suppose, on the other hand, you had no information or you were not concerned with simply verifying a number. Instead, your goal is to simply “measure the compliance rate.” In this case, you are measuring the value of the factor. In the first case you are performing a “one-tailed test” (i.e., you are concerned with only one side of a data distribution) while in the second you are performing a “two tailed test”. To determine which type of test is appropriate you should answer the two following sub-questions. First, do you have some information on the value of the factor you are measuring? Second, do you wish to verify whether or not the value is too high or too low? If you answer “yes” to both of these, then you are

³³ Most notably, you must develop an estimate of the variance of the sample before you can estimate sample size, which can be somewhat complicated. Sample sizes for proportions also require a variance estimate, but since proportions are bounded by 0 and 1, the variance is also bounded. Thus, as a worst-case scenario, you could use the maximum value for variance in drawing a sample for proportions.

conducting a one-tailed test. If you answer “no” to one or both of these questions, your situation is best characterized by a two-tailed test. *Most sample sizes can be drawn using two-tailed tests, which offer more conservative sample size estimates.* In other words, you can use a sample drawn using a two-tailed test to answer one-tailed questions, but you cannot do the reverse and remain statistically valid. If you are unsure about which type of test you are using, your best course of action is to assume a two-tailed test.

Question 3: With what level of error are you comfortable?

If you took a census of a population, you could be 100 percent certain of your results. In taking a sample, however, you are making an estimate of some parameter that you then extrapolate, with some uncertainty, to the total population. The purpose of sampling is to reduce the cost of measuring the parameter, while obtaining an estimate that is close to the population value. Thus, your estimate will involve some level of error. In using standard sampling procedures, however, you choose a level of error with which you are comfortable.

In sampling for a proportion, you are actually choosing a range interval around your final estimate that you are fairly certain contains the population value. As an example, suppose you chose an error of five percentage points and your resulting estimate of a compliance rate was 65 percent. This would mean that you are fairly certain the compliance rate for the population is between 60 and 70 percent (i.e., plus or minus five percentage points). One way to think of ‘error’ is that it defines the interval that you are comfortable saying the population parameter falls in. In the example, the interval is 10 percentage points (two times the error we chose because the error applies to both sides of the estimated value). If you had chosen a 10 percentage point error and obtained the same 65 percent compliance rate, then your extrapolation would be much less precise: (i.e., you are fairly certain that population compliance is between 55 percent and 75 percent). The more precise your estimate (smaller error), the larger the sample size. In fact, you may want to calculate sample sizes for more than one error level.

The selection of error level will depend on what you consider to be relevant or acceptable. Consider the above example where we used a 10 percentage point error. We are fairly certain that compliance is between 55 and 75 percent. This level of precision may be fine for some projects, but too imprecise for other. Additionally, if you are making a statistically valid comparison between two groups

(e.g., those receiving compliance assistance and those who do not) (see Question 6 below for more details), an error that is too large may result in not detecting a difference in compliance rates. For any project, you should determine which level of precision (or error) is acceptable to you given the goals of the project.

Question 4: How confident do you want to be in your estimate?

Question 3 used the phrase “fairly certain” to describe the confidence that the population compliance rate was within the interval. In estimating a sample size, you must also choose a level of confidence, defined in probability terms. Accepted research methods, however, provide very clear guidance on this aspect of choosing a sample size. In general, you should choose a level of confidence of 90 percent or higher; in reality, you need to consider only two levels of confidence: 90 and 95 percent.³⁴ If you had chosen 95 percent in the above compliance rate example (with the 5 percentage point error and 65 percent estimate), then you are 95 percent confident that the population compliance rate is between 60 and 70 percent. Note, however, that there is still a 5 percent chance that the population rate is not within that interval. Additionally, if you use a sampling frame that biases your results (see discussion above) or collect data in a manner that biases your results (see discussion below), then these statements on confidence and precision are meaningless. A sample size is only one aspect of choosing a valid random sample. You must perform all of the other aspects correctly to ensure that your statements on precision and confidence are valid.

Question 5: Are you concerned with making statistically valid statements about sub-groups of the population?

This is an important question and can have significant implications for your final sample size, and thus on the resources needed to administer your survey. As an example of this, you may be concerned with measuring compliance among auto repair shops, and you have divided them into two groups: independent shops and dealership-based ones. In this situation, you could (a) take a statistically valid measurement for auto repair shops as a whole or (b) take a statistically valid measurement for each subgroup individually. The second would imply the first, but requires significantly more resources. You

³⁴ You cannot choose 100 percent since you can only be 100 percent certain about your estimate if you perform a census. Mathematically, 100 percent confidence reduces the formula for sample size to an undefined answer.

may want to do the second if your decision to implement a program to either subset depended on the compliance rate for that group.

What often limits the use of taking valid measurement for sub-groups of a population is resource constraints. In fact, you may decide that making statistically valid statements about subgroups is not a priority after estimating a sample size for such a task. Making statistically valid statements about subgroups of the population amounts to making statements such as the ones above (e.g., 95 percent confident of the population value being within 5 percentage points on either side of the estimated sample value) for each subgroup. To do this, you would need to treat each subgroup separately and estimate a valid random sample size *for each group*. This could substantially increase the total sample needed for the survey project. You could still estimate parameters for subgroups of the population without drawing a valid random sample for those groups, but this will generally lead to weaker statements than if you had drawn a valid random sample³⁵. Note that estimating valid parameters for subgroups of the population is different than *comparing* two (or more) sub-groups to one another. This is discussed below in more detail.

Question 6: Is the ultimate purpose of your sampling to compare two (or more) groups to one another or to compare a single group to itself at different periods of time?

In many cases, you want to compare two samples. The most common situation in a compliance assistance framework would be to compare a characteristic (e.g., compliance rate) before and after a compliance assistance program is implemented. Another example would be where you want to compare two different groups. For example, you may want to compare compliance rates among those that have received compliance assistance versus those that have not. To make the *comparison* statistically valid, you cannot simply draw two valid random samples (one before and one after). You must incorporate the actual hypothesis test into your sample size calculation.³⁶ As shown below, the formula for a two-sample

³⁵ From a statistical standpoint, you would leave the confidence level as is and let the error level (confidence interval) increase. Thus, your resulting estimate would be less precise.

³⁶ A hypothesis test is a statistical comparison of an actual value (e.g., the proportion of facilities that said “yes” to a specific question) to a hypothesized value. The comparison takes into account that the actual value may have some variability associated with it (i.e., variance). The purpose of such a comparison is to determine if the actual value is significantly different than the hypothesized value, given the variability of the actual value. Sample size estimates are based on constructing a hypothesis test. Most basic statistics text books will contain a brief discussion of the connection between sample size estimates and hypothesis

test differs substantially from a one sample test. To estimate the sample size for a comparison, you need to ask Question 2 with the following changes:

What difference in the parameter (e.g., compliance rates) do you want to detect? In terms of compliance rates, a 2-point difference may not be significant from a policy perspective, but a 10-point difference may be significant. You will need to determine the number of units that represents a significant difference in terms of your programs or policies.

You can compare two drawn samples without using the two-sample comparison approach. Such a comparison, however, would not be statistically valid (see definition above) and would result in weaker statements regarding the comparison. Nevertheless, such a comparison may still be acceptable depending on the purpose of your evaluation.

If the primary goal of your performance measurement project is to make a valid comparison, then you should follow the two-sample approach. One form of valid comparison that is often done in a performance measurement context is the “before and after” comparison (e.g., comparing compliance in a sector before a compliance assistance program to compliance after a compliance assistance program has been set up.) To perform a statistically valid “before and after” analysis you must use the two-sample approach. If a comparison is not the primary goal, then taking two valid random samples and comparing them using a standard hypothesis test should suffice.

Formulas for Sample Sizes

You can develop sample sizes for estimating a proportion by answering the six questions above and applying those answers to the formulas in Figure 5-1. If you are making statistically valid comparisons between two groups, or comparing one group to itself at different times, then you should use the two-sample formulas (Question 6). Otherwise, you should apply the one sample formulas. You would also apply the one-sample formula to each subgroup of a population if you wanted to make valid statements about each subgroup (see Question 5).

Estimating a sample size is a two-step process. First, you calculate a standard formula for either the one-sample or two-sample scenario. Then, you adjust that sample size estimate using the adjustment formula given in Figure 5-1. The adjustment is necessary for purely statistical reasons, but from an implementation standpoint results in reducing the sample necessary for making a valid measurement.³⁷

EPA suggests that you estimate several sample sizes by varying confidence and precision (error). This will allow you to look at the trade-off between resources (i.e., larger samples require more resources) and the statistical properties of your results.

Figure 5-1. Formulas for Estimating Sample Sizes

Sampling Procedure	Initial Estimate of Sample Size	Final (Adjusted) Sample Size Estimate
One sample	$n_o = \frac{z^2 p(1-p)}{d^2}$	$n = \frac{n_o}{1 + \frac{n_o}{N}}$
Two sample	$n_o = \frac{z^2 [p_1(1-p_1) + p_2(1-p_2)]}{d^2}$	
<u>Definitions</u>		
n_o	The initial (raw) estimate for the sample size. In cases where the population is large, you can use this as the sample size estimate.	
n	The final sample size estimate, adjusted for population size. For the two-sample comparison, the estimate applies to both samples (i.e., you need to pull a random sample of n for <i>both samples</i>).	
N	The population size.	
p, p_1, p_2	The proportions that you are estimating, in numerical form (i.e., 50 percent is $p = 0.5$). In the two-sample case, the subscripts refer to the two different samples. To calculate sample sizes, you must use assumed values for these. If you have some prior information on the proportion, use that information. For example, if you have anecdotal evidence that compliance is about 60 percent, then use $p = 0.6$. In most cases, you will have no prior information, so you would use 0.5 as the default value for the proportions. Using the default value also produces to the most conservative estimate and hence the largest sample size.	

³⁷ In statistical terms, the initial estimate assumes an infinite population. This is rarely the case in sampling projects, so you need to adjust the sample size estimate to reflect that your population contains a finite number of entities.

<u>Definitions</u>	
z	A statistical parameter that corresponds to the confidence level. It is the standard normal score for a given level of confidence. The following criteria should be used to choose a z-score value (see Question 2 for discussion of two- and one-tailed tests): <i>Two-tailed test:</i> If you have chosen 90 percent confidence, then use $z = 1.645$. If you have chosen 95 percent confidence, then use $z = 1.96$. <i>One-tailed test:</i> If you have chosen 90 percent confidence, then use $z = 1.29$. If you have chosen 95 percent confidence, then use $z = 1.645$.
d	The error level, or difference between two samples, that you have chosen. As with the proportions above, use the numeric forms of the proportions (e.g., 5 percentage points is 0.05).

E. Drawing the Sample From the Population

After developing the sample size, you need to draw the random sample from the population. This section discusses your sampling approach and how you adjust for non-response in your sampling plan.

Sampling Approach

The sampling approach specifies how you choose entities to be in the sample. In other words, if you have a population of 100 entities and have estimated a sample size of 20, which 20 entities do you choose? Clearly, the 20 should be randomly chosen, but how do you ensure randomness? Also, what if you want to be sure that some members of each important group are included in the sample? Can you do that without sacrificing randomness?

In most cases, you can choose a simple random sample from your population. To do this, simply assign each entity a random number and sort by that random number.³⁸ Your sample will then be the first n entities, where n is your sample size.

Not all situations are this simple, however. For example, you may want to ensure that some members of specific subgroups are included in the sample, but you may not want to use a valid random sample for each subgroup due to resource constraints. In this situation, you can use a proportional

³⁸ Most spreadsheets contain a random number generator. Once you have assigned a random number to each entity, change the random number formula to a value since most spreadsheet random number generators recalculate each time the spreadsheet recalculates.

sampling framework which is generally referred to as stratification. To do this, divide the population list into the relevant subgroups and calculate the proportion that each represents. For example, we may have 100 entities, 75 use process A and 25 which use process B. Given the large number using A, you can select a sample of 20 without any process B entities. To ensure some representation of B entities, you could choose 75 percent of our sample from the A entities (i.e., 15 entities) and 25 percent from the B entities (5 entities). The key is to choose the sample randomly from each group. Thus, you follow the simple random sampling procedures described above for each group. Note, however, that you do not have a valid random sample for A entities or B entities *individually*, but a valid random sample for the population of A and B entities as a whole.

One final possibility is to use a cluster sampling procedure. Cluster sampling involves dividing the entities of the sampling population into “clusters,” randomly choosing those clusters as a first step in the sample draw, and then randomly choosing the entities within the clusters. Cluster sampling is a procedure that is best left to qualified statisticians, but understanding when it may be useful could be helpful.

Specifically, a cluster sample is useful when:

- # You cannot develop a sampling frame due to logistical or practical reasons. In this situation, use cluster sampling to develop the frame by first choosing clusters and then entities within the cluster. For example, if you were looking issues related construction sites in a certain state, a comprehensive list of those sites may be difficult or cumbersome to obtain. One method to locate those sites, and thus form a sample frame, would be to first divide the state into its counties and then randomly select counties. After that, you could obtain lists of construction permits from each county and then randomly pull sites from those lists.
- # Performing the survey on a given random sample would be expensive due to the geographic distribution of entities. For example, you may be conducting site visits that require substantial travel. You could reduce your costs if one trip included 2 to 3 site visits. In this situation, you could develop a cluster sample based on location to reduce costs and preserve the randomness of the sample.

Nonresponse Issues

So far, the discussion has assumed that entities will respond to your survey requests. In reality, this is rarely the case. Section IV of this guide discusses methods of reducing nonresponse. Although those methods are effective, completely eliminating nonresponse is almost impossible. If you anticipate

some nonresponse in your survey, adjust your sample size accordingly. The simplest way to do this is to inflate the sample size to cover any anticipated nonresponse. For example, if you anticipate that 50 percent of all potential entities will not respond, then double your sample size. You should also track who the nonresponders are, so that you can analyze any trends among nonresponders.³⁹ If you suspect that nonresponse will be a serious problem for your survey, however, you may want to consider using a qualified survey statistician.

³⁹ For example, if you have size information on the population and most of your nonresponders were “small” facilities, then your results may be biased since large facilities would be over-represented in the data.

SECTION VI: OECA'S EXPERIENCE WITH THE DILLMAN METHOD AND STATISTICALLY VALID SURVEYS

In 2001, EPA's Office of Compliance (OC) tested methods for collecting outcome data from compliance assistance efforts. The study compared two survey methodologies: mailed surveys using the Dillman "total design method" and surveys conducted as on-site observations. The purpose of this analysis was to test the hypothesis that the results of both data collection methods would be the same. For this effort, EPA analyzed sectors in different stages of compliance assistance activity: 1) a sector for which EPA is beginning a compliance assistance effort (marinas); and 2) a sector for which OC, other EPA offices, and states have conducted several compliance assistance activities (metal finishing). EPA also used the survey results to evaluate the impact of current compliance assistance activities and to determine sector specific compliance assistance issues. However, this section focuses on comparing the two data collection approaches.

A. Background and Purpose of the Study

EPA studied metal finishers and marinas to compare data gathering approaches and to evaluate the effectiveness of the Dillman method for collecting compliance assistance evaluation data. Due to resource limitations, EPA focused the metal finishing study on EPA Regions 1 and 5. EPA focused the marinas study on EPA Region 1 since this region was beginning a marinas compliance assistance initiative. Table 6-1 summarizes background information for each study, including the target population, the survey scope, consultations conducted for survey development, and the source of the list used to develop each sample frame. EPA conducted the following steps for each sector:

- # Develop the survey instruments;
- # Determine the number of survey respondents required for each survey (see Section 6.B);
- # Develop the list of recipients for the mailed and on-site observations (see Section 6.B);
- # Use the Dillman method for the mailed survey (see Section 6.C);
- # Arrange and conduct the on-site observations (see Section 6.C); and
- # Analyze the results (see Section 6.D).

Although the metal finishing and marinas survey projects both consist of a mailed survey and on-site observations, the results from the two sectors are not directly comparable. The goals of the two projects were vastly different. The metal finishing survey project was designed to evaluate the relative

usefulness of mailed surveys and on-site observations. The metal finishing survey focused on awareness of regulations and general indicators of environmental performance. The marinas survey project, on the other hand, was initially designed to collect baseline compliance information through on-site visits. OECA added a mailed survey to this project to evaluate the relative effectiveness of each survey mode. Given the differing original goals of the two projects, comparing the results is not valid. Additionally, the two sectors are vastly different with respect to their experience with environmental regulations. The metal finishing sector has a great deal of experience with environmental regulations, but, in contrast, the marinas sector has very little regulatory experience. This difference may affect the level of regulatory awareness that each sector has.

In developing the survey instruments, EPA consulted with a variety of industry representatives and compliance assistance providers. Table 6-1 provides more information on the groups consulted for this effort. For each sector, EPA used the same survey instrument for both the mailed survey and the on-site observation (i.e., for the metal finishing study, the same survey instrument was used for both the mailed and on-site observations). Figure 6-1 presents the metal finishing survey. Figure 6-2 presents the marina survey. These surveys are included at the end of the section.

Table 6-1. Background Information on EPA Data Collection Study

	Metal Finishing Sector	
Goals	The goals for the metal finishing sector survey were to (1) collect information through two survey modes and compare the costs and results of the two survey modes and (2) collect basic information on the metal finishing sector that could be used in developing compliance assistance programs.	The goal of the marinas survey was to collect baseline compliance information using both self-reported information and an on-site visit program as part of an effort to measure the results of Region 1's environmental assistance program. Self-reported data were collected through a mailed survey to compare to the on-site data collection.
Target Population	Metal Finishers in EPA Regions 1 and 5. These regions have conducted extensive compliance assistance for metal finishers.	Marinas in EPA Region 1. This region was initiating a compliance assistance program for marinas.

	Metal Finishing Sector	Marinas Sector
Survey Instrument Consultation	In developing the metal finishing survey instrument, EPA/OC consulted with: <ul style="list-style-type: none"> • Members of the metal finishing industry; • Metal finishing trade associations; • State compliance assistance staff; and • Other EPA offices. 	In developing the marinas survey instrument, EPA Region 1 conducted stakeholder meetings to consult with: <ol style="list-style-type: none"> 1. Marina owners; 2. Four state marina trade association presidents; 3. State compliance assistance providers; 4. State Coastal Zone Management program offices; and 5. Small Business Development Center staff members.
Sampling List Data Source(s)	The Harris Directory - A database with information on more than 700,000 U.S. businesses, covering the manufacturing sector. Information includes company name, location, size, and industry type (determined by SIC code). EPA purchased a list of all facilities in SIC codes 3471 and 3479 in the EPA Region 1 and 5 states.	State Coastal Zone Management program offices provided much of the marina list. Additional information for inland sites was obtained from tourism and boating guides and state boating maps.

B. Methodology

Estimating Sample Size

To estimate a sample size for this project, EPA followed the methods discussed in Section V of this guide. Table 6-2 provides EPA's answers to the six questions posed in Section V. Based on the answers to the six questions, EPA applied the sample size formula for a two-tailed, one-sample procedure found in Figure 5-1 along with the suggested adjustment to the initial estimate. EPA calculated that a sample size of 92 to 94 facilities was necessary for the metal finishing sector and that a sample size of 70 facilities was necessary for the marinas sector.^{40,41}

⁴⁰ The ranges stem from the use of ranges for the population size in the adjustment formula of Figure 5-1. For metal finishing, a population range of 2,000 and 4,000 facilities was used. For Marinas, a population of 1,200 facilities was used.

⁴¹ During the sample facility selection phase for marinas, however, one state exited from the study. As discussed later in this section, both the metal finishing and the marinas studies involved stratifying the sample size by state. To account for the exited state, the sample size for marinas was reduced by 30 facilities which was the number that the state accounted for in the stratification process.

Table 6-2. Factors Contributing to Sample Size Estimate

Question	
1: Proportion or Average Value?	EPA's goal (for metal finishing) was to measure the percentage of facilities that had received compliance assistance. This implies that the formulas for proportional sampling (found in Section V) can be used. Additionally, EPA had no information on this proportion to construct an initial estimate. Therefore, for purposes of estimating the sample size, EPA assumed $p = 0.5$ in the Figure 5-1 equations.

Question	Metal Finishing and Marinas Survey
2: One- or Two-Tailed?	EPA designed this analysis as a two-tailed test to learn the compliance rate rather than the percentage above an assumed rate. This implies using the z values for a two-tailed test found in Figure 5-1 (see question 4 below for an exact value).
3: Level or error (precision)?	EPA chose a 10 percentage point level of precision, or $d = 0.10$ in the Figure 5-1 formulas.
4: Confidence?	EPA chose a 95 percent level of confidence. Combined with question #2, this implies using a z value of 1.96 in the Figure 5-1 formulas.
5: Statistically valid stratification?	EPA did not stratify the sample for purposes of making statistically valid statements about subgroups of the population. EPA did, however, stratify to ensure representation of the different sizes, states, and types of facilities. Therefore, the sample will represent the population as a whole and not a subgroup.
6: Statistically valid comparisons?	EPA's goal was <i>not</i> to make statistically valid comparisons. This implies using the one-sample formula.

Stratification

Metal Finishing

EPA stratified the random sample by both state and size (defined by number of employees) for both the mail survey and the on-site observation. This was done to ensure that each of the states in EPA Region 1 (CT, ME, MA, NH, RI, VT) and EPA Region 5 (IL, IN, MI, MN, OH, and WI) were represented in the surveys and that both small and large facilities were included in the sample.⁴² For this project, EPA defined any metal finisher with less than 10 employees as a “small” facility and all other facilities as “large.”

EPA stratified first by state by using information contained in the sampling frame to determine the percentages that each state represented in the population. EPA applied these percentages to the sample size estimate to determine the number in the sample from each state. Next, EPA used information from the sampling frame to determine the percentage of the population *within each state* that employed less than 10 employees. EPA applied these percentages to the sample size for each state to determine the number of small and large facilities to draw from each state.

⁴² Although EPA intended to ensure that all states and sizes were represented, EPA did not intend to make statistically valid measurements of these subgroups (i.e., each state and large and small facilities).

Marinas

EPA stratified the sample for marinas by both state and location (inland versus coastal) for both the mailed and the on-site observations. As with the metal finishing surveys, the stratification was done to ensure that each state and both location types were adequately included in the sample taken.

EPA stratified first by state by using information contained in the sampling frame to determine the percentages that each state represented in the population. EPA applied these percentages to the sample size estimate to determine the number in the sample from each state. EPA then used the percentage in each state that were inland and coastal marinas in the sampling frame to estimate the number of each marina type in each state's sample.

Sample Selection and Nonresponse Replacement

Metal Finishing

EPA conducted the surveys on this sector as a blind voluntary random sample (i.e., the facility identity is unknown to the Agency and facilities participated voluntarily). To draw the random sample, EPA first divided the population into separate lists for each state and then within each state by size class (i.e., those with less than 10 employees and those with 10 or more employees). EPA assigned each facility a random number (using Microsoft Excel's random number generator) and then sorted each list based on random number.

For the metal finishing sector, EPA anticipated at least a 50% response rate on the mailed surveys using a four step contact process. To account for this, EPA doubled the sample size estimate for mailed surveys. As a result, EPA actually mailed 202 surveys to facilities, more than twice the initial sample size estimate.

EPA anticipated that each facility visited in the site observation would result in a completed survey. In the event that some facilities did not agree to participate in the site visit program, EPA continued to pull facilities from the sorted list of facilities. A number of facilities refused to participate

(see Section 6D, Results below). Project schedules forced EPA to discontinue the site visit phase after visiting 67 in-scope facilities.

To ensure that no facility appeared in both the mailed and visited samples, EPA drew both samples simultaneously. The first $2x$ facilities in each sorted list of state/size combinations constituted the random sample for the mailed survey, where x was the original sample size estimate for each state and size combination. The next x facilities were the initial set of contacts for the site visit observation. If a member of the site visit group refused to participate or was found to be out of scope, EPA contacted the next facility on the list. In total, EPA drew a random sample of more than three times the desired range, randomly assigning two-thirds of the sample to the mail survey and one-third to the site visits.

Marinas

EPA conducted the marinas survey as a voluntary random sample. EPA selected the marinas for the on-site observations first, and then set up the required number of visits. Facilities that were not scheduled for an on-site observation were then used as the frame for the mailed survey sample. Thus, it was not possible for any facility to be in both samples.

EPA used a random selection process to develop the list of marinas to contact for the on-site observations. Specifically, EPA developed a final list for each state and then chose every tenth marina from the list until the required number of marinas was drawn for each state. EPA then reviewed the list to determine how many were inland versus coastal. If the required number of inland marinas was not attained for each state, EPA replaced the final coastal ones that were drawn with inland marinas that were also drawn randomly.

As with the metal finishing survey, some facilities declined on-site observation requests by EPA or were not in scope for the survey effort. To replace these facilities, EPA located the facility that had declined or was not in scope in the sample frame list and then selected the next facility in the list. If that replacement facility declined or was not in scope, EPA selected the one directly before the original selected facility. Additional replacement facilities were chosen in this manner, moving progressively further from the original selected facility and alternating between succeeding and preceding facilities.

EPA anticipated at least a 50 percent nonresponse rate for the mailed survey and in response more than doubled the number of surveys mailed. In total, EPA mailed 236 surveys to marinas. To choose the random sample for marinas, EPA followed an identical process as that for the metal finishing mailed survey.

C. Implementation

Mailed Surveys

In accordance with the Dillman method, EPA sent each facility a prenotification letter. This letter was signed by the appropriate regional authority and copied onto EPA regional letterhead. The letter introduced the project, described the purpose of the study, notified the facility of the imminent arrival of a mailed survey, and informed them that participation was voluntary. Appendix E contains a copy of the metal finishing prenotification letters.

The survey and a signed, copied, cover letter followed approximately two weeks later. The cover letter reiterated much of the prenotification information. It also provided a toll-free number to call with any questions regarding the survey or the program itself. Although the cover letter was copied onto EPA letterhead, the survey was mailed in a 9x12 manilla envelope showing the EPA contractor's return address information, rather than EPA's. This was done to emphasize the fact that this was a blind (to EPA) study. Each survey was printed with a randomly chosen site identification number to help the contractor evaluate if the sample frames (i.e., states, sizes) were equally represented in the responses. A stamped self-addressed return envelope was also included with the survey to help increase the return rates. Appendix E contains a copy of the metal finishing cover letter.

Once the facilities had been given adequate time to complete and return the surveys (2-3 weeks) a reminder postcard was sent to all facilities that had not yet responded. Each postcard was oversized (8.5" × 5") and printed on blue cardstock to make it more visible to the recipient. The post card text reminded facilities of the project and survey, thanked them if they had already participated, and set a deadline for returning the survey. Appendix E contains a copy of the metal finishing reminder postcard.

Finally, facilities that did not return the survey two weeks after receiving the postcard were sent a second survey. The re-mail package contained a modified cover letter. In addition to all of the previous information, this letter stated that the facility had already received a survey but had not yet participated. This package also contained a stamped, self-addressed return envelope. Appendix E contains a copy of the metal finishing re-mail cover letter.

As facilities returned their surveys, EPA's contractor logged them into a database. If any of the materials were returned as undeliverable, the facility was removed from the remaining mailing lists. These facilities were not considered as part of the "returned" surveys discussed in the "Results" section.

Site Visit Observations

As stated above for both the metal finishing and the marinas on-site observations, EPA selected a random set of facilities to contact for site visits. For the metal finishing sector, these facilities received a prenotification letter very similar to that of the survey recipients but this one notified of an upcoming phone call to schedule a site visit. Within two weeks the facilities were called to schedule an appointment. If a facility chose not to participate, could not be reached, or was not currently performing metal finishing activities, then a replacement facility was taken from the list. Replacement facilities, which were randomly selected at the beginning of the process, were chosen based on both size and state. For example, if a small site in Connecticut chose not to participate, then the next small, Connecticut site was taken from the top of the replacement list. This process was repeated until either the total desired number of site visits were scheduled or the project schedule/budget was expended.

For the marinas sector, once a set of facilities was selected, each facility was contacted to determine eligibility (i.e., in scope or not) and willingness to be involved in an on-site observation. Facilities that were either out-of-scope or declined a visit were replaced using the process described in the previous section. To encourage participation, the marina owner was told that assistance would be offered during the visit and it was clarified that the visit was not an official enforcement inspection.

D. Results

This section compares responses to the two survey modes (e.g., response rates). We compare the two samples with respect to their general characteristics, the types of compliance assistance received, and the understanding of regulatory requirements. These three comparisons support the inference that the samples are similar in nature (i.e., the respondents for each survey mode had similar characteristics). Finally, we compare the performance towards key environmental regulations.

Response Comparison

Several aspects of response rates are relevant to compare for the two samples. We begin by discussing how we calculated the overall response rate for each survey mode. Next, we focus on the details of response for the mailed surveys and then the details of response for the site visit observations. Finally, we compare the resource requirements to implement each survey mode on a per response basis.

Overall Response Rates

To calculate response rates, EPA followed guidance from both the Council of Survey Research Organizations (CASRO) and Lohr (1999).⁴³ Both sources suggest adjusting for out-of-scope entities among unreturned surveys in calculating response rates. In cases where we can be sure that *all* members of the sampling frame are in-scope, dividing the number of returned surveys by the total number of surveys sent will provide an accurate response rate. In most survey situations, including this one, we cannot be sure that all members of the sampling frame are also in-scope. Thus, some entities may not return a survey since they are out of scope. Survey researchers tend to agree that out-of-scope facilities that do not respond should not be counted as non-responders when calculating response rates. Therefore, EPA adjusted for this by estimating the number of out-of-scope facilities that did not respond.⁴⁴

⁴³ The CASRO guidance was the result of a task force developed by CASRO to provide definitive guidance to survey researchers on how to calculate response rates.

⁴⁴ In the formulas that follow, this is done in the denominator of each response rate. The first term of each denominator multiplies the percentage that are in-scope among the returned surveys (e.g., CIS/TR in the mailed survey response rate) by the number of unreturned surveys (NR). Thus, the first term provides an estimate of the total in-scope non-responders by using information from the returned surveys.

We used the following formulas to calculate response rates (*RR*):

<p>Mailed Surveys</p> $RR = \frac{CIS}{\frac{CIS}{TR} \times NR + CIS}$ <p>RR = Response Rate CIS = Completed in-scope surveys TR = Total Returned NR = Number Not Returned</p>	<p>Site Visits</p> $RR = \frac{V}{\frac{IS}{IS + OS} \times U + IS}$ <p>IS = In-scope facilities = V + M + R V = Number of visits. M = Facilities that agreed to receive visit, but we were unable to visit. R = In-scope facilities that refused visits. U = Unknown (unreachable) OS = Out-of-scope facilities.</p>
---	---

In developing these response rate formulas, EPA concentrated on developing estimates that could be compared across the two survey modes. Table 6-3 presents the estimated response rates for both sectors and both survey modes.

Table 6-3. Response Rates

Mode	Metal Finishing	Marinas
Mailed	43.6%	51.3%
Site Visits	40.7%	94.6%

For the metal finishing sector, both evaluation modes resulted in response rates less than 50 percent. Additionally, the mailed survey resulted in a slightly higher response rate (by 2.9 percentage points) than the site visits. Normally one would expect a significantly higher response rate for the site visits; however, there is a high level of distrust towards the regulatory agencies and therefore a degree of hesitancy towards voluntarily allowing them on site. In addition, the mailed survey response rate is much higher than the 17 percent that OC has experienced in the past. This success is mostly attributed to the Dillman method. The Marinas’ surveys both had response rates over 50 percent, albeit the mailed response rate was only 51.3 percent. The site visit response rate, however, was close to 95 percent, indicating a high degree of success in getting facilities to agree to participate. This may be indicative of the sector which in the midst of a contentious regulatory development program (metal finishing) an one that has no negative history with EPA (marinas), as well as other differences between the programs.

Details on the Response Process for Both Mailed Surveys and Site Visit Observations

Table 6-4 presents information on the return process for the mailed surveys for both the metal finishing and the marinas sectors. As shown in Table 6-4, more than half of all responses resulted from the initial mailing for both surveys. Additionally, the initial mailing and the postcard generated more than 70 percent of all responses in both sectors (80.9 percent for metal finishing and 71.8 percent for marinas).

Comparing the total mailed to total in-scope returns, we see that the metal finishing survey received one in-scope response for every three mailed surveys, while the marinas sector received one in-scope response for every two mailed surveys. This is likely due to the data sources used to identify facilities in each sector. Since the metal finishing study included a broader geographic range than the marinas sector, EPA used a national database (the Harris Directory) to identify recipients. This database likely included more out-of-scope facilities than the more local databases used for the marinas survey. Additionally, more refinement and cross-checking was done for the marinas sampling frame due to the lack of one comprehensive source for a sampling frame as in the metal finishing surveys.

Table 6-4. Mailed Survey Returns

Survey Status	Metal Finishing	Marinas
Returned After Initial Mailing	52 (76.5%) [a]	71 (60.7%) [a]
Returned After Reminder Postcard	3 (4.4%) [a]	13 (11.1%) [a]
Returned After Second Mailing	13 (19.1%) [a]	33 (28.2%) [a]
Total In-Scope Returns	68	117
Total Mailed	202	236

[a] As a percentage of total in-scope returns.

Table 6-5 presents similar information for the site visits. For the metal finishing sector, a little more than half of all site visits (53.7 percent) came from the original sample. Additionally, comparing the total number of visits to the total number of contacts, it was necessary to contact 4.25 facilities to obtain each site visit in the metal finishing sector. Details on originals versus replacements were not available for

the marinas on-site observations. Comparing total contacts to number of visits, however, the marinas on-site observation project needed only 2.1 contacts to generate the required number of in-scope visits. The discrepancy between the number of contacts per in-scope visits may be associated with degree of prior EPA attention for each sector. That is, the marinas sector has received much less attention in the past from EPA compared to the metal finishing sector.

Table 6-5. Distribution of Responses for Site Visits

Contact Stage/Status	Metal Finishing	Marinas
Visits – Original Sample	36 (53.7%) [a]	NA [b]
Visited – Replacement Sample	31 (46.3%) [a]	NA [b]
Total Visits	67	70
Total Contacted Sites	285	146

[a] As a percentage of total visits.

[b] This information was not available for the marinas on-site observations.

Budget–Cost Per Response

Table 6-6 presents information on the cost per in-scope response for both sectors and both survey modes. Both the hours and the nonlabor costs included in Table 6-6 only reflect those items that are different between the two survey modes (e.g., the hours required for developing the survey instrument and mailing list and the costs for the Harris Directory are not included since these would be required for each survey mode). Although not included in the costs and hours below, EPA has found that defining the target population, developing the survey instrument, and acquiring a useable recipient list can be the most time consuming portions of the project.

When evaluating the numbers presented in Table 6-6, consider two important factors: 1) the limits on the types of information that can be collected through a mailed survey, and 2) the value of additional anecdotal information that can be collected through on-site observations. As discussed later in this section, mailed surveys are useful for obtaining general information on awareness of environmental regulations or information on trends in environmental performance. However, mailed surveys may not yield accurate data on actual compliance rates; this type of information can be collected during on-site observations. Also, during on-site observations, facility personnel may become comfortable enough to

openly discuss compliance concerns, root causes of non-compliance, and compliance assistance preferences. This type of anecdotal information is helpful in designing compliance assistance programs. For the metal finishing sector, each in-scope on-site observation required an additional 9.9 labor hours and approximately \$561 in costs. For the marinas, each in-scope on-site observation required an additional 5.3 labor hours and approximately \$274 in costs. Much of these differences can be accounted for by additional travel time and money required by the multi-regional scope of the metal finishing project.

Table 6-6. Cost Per In-Scope Response

Category	Metal Finishing	Marinas
<i>Site Visits</i>		
Labor Hours Per In-Scope Response	11.0 [a]	6.4 [g]
Labor Cost Per In-Scope Response	\$435 [b]	\$253 [b]
Nonlabor Costs Per In-Scope Response	\$177 [c]	\$72 [h]
<i>Total Cost Per In-Scope Response</i>	<i>\$612 [d]</i>	<i>\$325 [d]</i>
<i>Mailed Surveys</i>		
Labor Hours Per In-Scope Response	1.1 [e]	
Labor Cost Per In-Scope Response	\$43 [b]	
Nonlabor Costs Per In-Scope Response	\$8 [f]	
<i>Total Cost Per In-Scope Response</i>	<i>\$51 [d]</i>	

[a] Hours include training, phone calls to set up visits, travel time, on-site interview time.

[b] Calculated by multiplying the hours by a loaded hourly rate of \$39.51. The loaded hourly rate reflects use of a GS-13, Step 1 employee (Washington, DC area pay scale) in 2001. The annual salary for that employee is \$63,211, resulting in a hourly rate of \$30.39. We added a 30 percent mark-up to this hourly rate to get the loaded rate of \$39.51.

[c] Costs include only those costs unique to site visits, such as travel-related costs (airfare, hotel, car rental, food) for site visits and long-distance phone charge charges.

[d] This is the sum of labor and non-labor costs.

[e] Hours include coordination with Regions, copying, filling envelopes, and logging survey returns.

[f] Costs include only those costs unique to mailed surveys, such as copying, postage, envelopes, labels, and printing postcards.

[g] Hours include travel time, training, phone calls, and designing a visit protocol.

[h] Costs include travel-related costs associated with hotel, car rental, food, and incidentals.

Comparing Characteristics and Survey Responses*Metal Finishing*

Tables 6-7 through 6-9 summarize characteristics of the two samples for the metal finishing sector. Each characteristic was represented by a yes/no question in the survey. This information compares the two samples and determines how similar they are to one another. If there are few differences in these characteristics, then the two survey modes have essentially drawn comparable samples, and differences may be attributable to the survey mode. A large number of significant differences would indicate that respondents to one survey mode are different than the respondents to the other mode.

To formally compare the two samples, EPA performed statistical hypothesis tests. These tests looked for statistically significant differences between the two samples in terms of the percentage that said “yes” to the question. That is, did one sample have a significantly larger percentage that said “yes” to the question than the other sample. The results of these comparisons can be summarized as follows:

- # **General Characteristics (Table 6-7)**—For the most part, the two samples are identical in terms of general characteristics. There are, however, two significant differences. A larger proportion of the visited facilities generate both metal-bearing wastewater and hazardous waste compared to the mail survey facilities. This is not surprising for the following reasons. If a facility generates cyanide- or chromium-bearing wastewater, it usually also generates metal-bearing wastewater. These facilities are typically aware of the constituents in their wastewater and their regulatory requirements. However, if a facility generates only metal-bearing wastewater (e.g., they may only be performing barrel finishing or cleaning operations prior to dry surface finishing operations), it is possible that they would not consider their wastewater “metal-bearing” when completing the survey. The site-visit personnel, however, would classify this as metal-bearing wastewater because of expected metals concentrations. With respect to the hazardous waste generation, several site-visit personnel identified facilities that did not realize that they were generating hazardous waste even though they were aware of hazardous waste requirements; therefore, it is not surprising that the mailed surveys indicated fewer hazardous waste generators than the site visits.

- # **Compliance Assistance (Table 6-8)**—The types of compliance assistance received by both samples is generally similar. However, a larger proportion of visited facilities acknowledged receiving wastewater compliance assistance compared to the mail survey facilities.

- # **Awareness (Table 6-9)**—There were no significant differences between the two samples in terms of facilities' understanding of regulatory requirements.

Based on the few significant differences, the two metal finishing samples appear to be similar.

Table 6-7. General Characteristics of the Two Samples in the Metal Finishing Sector

Characteristic	Percent Answering "Yes"		Statistically Significant Difference
	Mailed Survey (n = 68)	Site Visits (n = 67)	
<i>Wastewater</i>			
Generate Chromium-Bearing Wastewater	36.8%	37.3%	No
Generate Cyanide-Bearing Wastewater	23.5%	20.9%	No
Generate Metal-Bearing Wastewater	47.1%	64.2%	Yes
Have a Discharge Permit	60.3%	56.7%	No
<i>Hazardous Waste</i>			
Generate Hazardous Waste	48.5%	62.7%	Yes
<i>Air Emissions</i>			
Perform Chromium Electroplating	22.1%	19.4%	No
Have a Solvent Cleaning Machine	19.1%	19.4%	No
<i>Toxic Release Inventory</i>			
TRI Exempt	29.4%	34.3%	No

Table 6-8. Types of Compliance Assistance Received For the Two Samples in the Metal Finishing Sector

Compliance Assistance Received	Percent Answering "Yes"		
	Mailed Survey	Site Visits	
Wastewater	17.6% (n = 43)	32.8% (n = 47)	Yes
Hazardous Waste	22.1% (n = 33)	32.8% (n = 42)	No
Air Emissions—Chromium Electroplating	8.8% (n = 15)	6.0% (n = 13)	No
Air Emissions—Solvent Cleaning Machine	5.9% (n = 13)	7.5% (n = 13)	No
TRI Reporting	20.6% (n = 68)	25.4% (n = 67)	No

National Metal Refinishing Resource Center web site	13.2% (n = 68)	9.0% (n = 67)	No
---	-------------------	------------------	----

Table 6-9. Awareness of Regulatory Requirements for the Two Samples in the Metal Finishing Sector

Awareness	Percent Answering "Yes"		Statistically Significant Difference
	Mailed Survey	Site Visits	
Hazardous Waste Labeling Requirements	97.0% (n = 33)	95.2% (n = 42)	No
Hazardous Waste Manifest Records Requirements	97.0% (n = 33)	95.2% (n = 42)	No
Air Emission Control Requirements for Chromium Electroplating Tanks	100.0% (n = 15)	100.0% (n = 13)	No
Air Emission Standards for Solvent Cleaning Machines	61.5% (n = 13)	69.2% (n = 13)	No
TRI Form R	61.8% (n = 68)	53.7% (n = 67)	No
National Metal Finishing Strategic Goals Program	23.5% (n = 68)	26.9% (n = 67)	No

Marinas

Table 6-10 summarizes the characteristics of the marinas in both the mailed and on-site observations. As with the information for the metal finishing surveys, EPA performed statistical tests to determine if significant differences existed between the samples in each mode. Unlike the metal finishing survey, however, a number of the questions we assess in Table 6-10 are quantitative in nature rather than yes/no questions. For the quantitative questions, we compare the average value between the two surveys.

We compared eight characteristics between the two samples and in four cases there was a significant difference. Most notably, the facilities contacted through the mailed surveys had more full-time employees than the on-site observations, but the on-site observation facilities had a larger number of boats (total boats moored, docked, and stored on land) than the mailed survey marinas. Additionally, of those marinas where maintenance and repair operations take place, a significantly larger proportion of on-site observations had primarily customer do-it-yourself maintenance and repair operations compared to the mailed survey sites. These three significant differences (number of full-time employees, total number of

boats, and primarily customer do-it-yourself maintenance and repair) make it difficult to conclude that the two samples are similar.

Table 6-10. Characteristics of Marinas in Both Samples

Characteristic	Mailed Survey	On-Site Observations	Statistically Significant Difference
<i>Number of Employees</i>			
Average number of full-time employees	13.6 (n = 117)	8.8 (n = 70)	Yes
Average number of part-time employees	3.7 (n = 117)	4.3 (n = 70)	No
<i>Number of Boats</i>			
Average number of moored boats	31.3 (n = 117)	27.1 (n = 70)	No
Average number of docked boats	70.7 (n = 117)	96.5 (n = 70)	Yes
Average number of boats stored on land	73.8 (n = 117)	94.9 (n = 70)	No
Average number of boats that are moored, docked, and stored on land at the facility	175.9 (n = 117)	218.5 (n = 70)	Yes
<i>Maintenance and Repair Operations</i>			
Percentage that indicated maintenance and repair (M/R) operations are performed at the marina	90.4% (n = 114)	94.2% (n = 69)	No
Of those that indicated M/R operations are performed, the percentage the indicated the M/R operations are primarily customer "do-it-yourself" operations.	16.2% (n = 105)	40.6% (n = 69)	Yes

Comparison of Performance Towards Key Environmental Regulations

Metal Finishing

Table 6-11 compares the results of the on-site observations to the mailed surveys for a number of compliance-related questions. The purpose is to compare the environmental performance reported in the mailed surveys to that observed on-site, given the assumption that the on-site observations would generate accurate information. We can then evaluate the correlation between the mailed survey and observed

information to assess if the mailed survey respondents biased their responses towards better environmental performance.

In total, we compared answers to 15 questions that specifically requested information on environmental performance. Table 6-12 summarizes the numbers and percentages of questions where mailed respondents had (a) better reported performance, (b) statistically significant better reported performance, and (c) a reported performance five percentage points or more higher than on-site observations (considered to be numerically significant). For eight of the 15 questions, the mailed responders indicated better performance towards key environmental regulations. On the remaining seven, the performance was either the same between the two modes or the on-site responses had better performance. For two questions, the compliance rate for mailed respondents was larger by a statistically significant margin. In six of the questions the mailed respondents had a numerically significant (five point or larger) performance rate. Finally, there was one specific trend: for all performance-related questions for the air section of the survey, the mailed surveys indicated a better performance.

Based on these results, we can conclude that in general the mailed survey responses do not provide a biased estimate of performance towards key environmental regulations, but there may be some concerns for the air-related questions. One of the two cases where mail responders indicated significantly better performance was an air-related question. Combined with the fact that mail responders indicated better performance for all air-related questions, this may indicate some concern about the accuracy of the mail responders for these questions.

There is one caveat to this analysis, however. The results are all based on small numbers of respondents for each question. Specifically, the on-site questions had an average of 22 respondents with a range of 6 to 38 respondents and the mailed survey had an average of 24 respondents with a range of 7 to 41 respondents.⁴⁵ If we had more respondents, we may have found more significant differences. Nevertheless, we expect that our results are representative enough to conclude that, except for the air questions, the mail responders accurately depicted their performance.

⁴⁵ These numbers differ from the total response for the surveys and from question to question, since not all questions applied to all respondents.

Table 6-11. Compliance Comparison for Metal Finishing Sector

Compliance Category	Mailed Percentage	On-site observation Percentage	Better Performance	
Wastewater Questions				
Have Chrome Permits and Chrome Treatment	87.0% (n = 23)	70.0% (n = 20)	Mailed	Yes
Have Cyanide Permits and Cyanide Treatment	93.3% (n = 15)	90.0% (n = 10)	Mailed	No
Have Metal-Bearing Permits and Metal-Bearing Treatment	93.5% (n = 31)	100.0% (n = 34)	On-Site	Yes
Conduct Monitoring as Required by Permit	92.7% (n = 41)	84.2% (n = 38)	Mailed	No
Conduct Reporting as Required by Permit	87.8% (n = 41)	89.5% (n = 38)	On-Site	No
Conduct Sampling as Required by Permit	85.4% (n = 41)	89.5% (n = 38)	On-Site	No
Hazardous Waste Questions				
Comply with Time Limited Storage Requirements	100.0% (n = 22)	95.5% (n = 22)	Mailed	No
Comply with Labeling Requirements	77.3% (n = 22)	100.0% (n = 22)	On-Site	Yes
Conduct Storage Area Inspections	100.0% (n = 22)	100.0% (n = 22)	Same	No
Conduct Inspections with the Required Frequency	86.4% (n = 22)	95.5% (n = 22)	On-Site	No
Have Contingency Plans	86.4% (n = 22)	86.4% (n = 22)	Same	No
Air Questions				
Have Emission Controls (Relating to Chromium) on Every Chromium Tank	93.3% (n = 15)	84.6% (n = 13)	Mailed	No
Have one of the Recommended Control Technologies	100.0% (n = 15)	92.3% (n = 13)	Mailed	No
Both Use and Monitor the Control Technology (Both are Required)	86.7% (n = 15)	53.8% (n = 13)	Mailed	Yes

Compliance Category	Mailed Percentage	On-site observation Percentage	Better Performance	
Have a Solvent Cleaning Machine with 5% or Greater by Weight of a Regulated Solvent and Know that the Solvent Cleaning Standards Apply	71.4% (n = 7)	66.7% (n = 6)	Mailed	No

Table 6-12. Summary of Compliance Question Results for Metal Finishing – Comparison of Mailed Surveys and On-Site Observation Compliance Rates

Category	Number (Total Compared = 15)	Percentage of Compliance Questions
Questions where mailed surveys participants had a higher performance than on-site observations	8	53%
Questions where mailed survey participants had a <i>statistically significantly</i> higher performance than on-site observations	2	13%
Questions where mailed survey participants had a higher performance <i>five percentage points or more</i> than on-site observations	6	40%

Marinas

Tables 6-13 to 6-16 compare the compliance performance between the marinas that participated in the two surveys. Table 6-17 summarizes the numbers and percentages of questions where mailed respondents had (a) higher reported compliance, (b) statistically significant higher reported compliance, and (c) a compliance rate five percentage points higher. As with the metal finishing analysis above, the purpose is to compare the mailed survey compliance rates to the on-site observations compliance rates under the assumption that the on-site observations would generate an accurate compliance rate. For each compliance-related question we perform a statistical test to determine if there was a significant difference between the two samples.

In total, we compared 38 questions and found that in 29 cases (76 percent of the questions) the mailed responders indicated higher compliance. For 23 of those questions, the mailed responders indicated a statistically significant higher compliance rate and in 22 cases the mailed responders had compliance rates five percentage points or higher. Among the nine cases where the on-site observations resulted in higher compliance, only four were significant. Furthermore, for all of the hazardous waste questions (10

questions total; see Table 6-13) and all of the hazardous materials questions (3 questions total; see Table 6-15) the mailed responders indicated significantly higher compliance.

This analysis shows that the mailed responders indicated significantly higher compliance than the on-site observations found. Thus, compliance-related questions from the mailed survey may not accurately reflect compliance. As noted above in comparing the characteristics of the samples, however, it appears that the two samples may be different, which may result in some discrepancies between the two modes. Nevertheless, the widespread differences between *reported* compliance by the mailed responders and *observed* compliance during the site visits should call into question the validity of obtaining reliable compliance information through a mailed survey.

Table 6-13. Hazardous Waste Compliance for Marinas

Performance Characteristic	Mailed Percentage	On-Site Visit Percentage	Better Performance	
Manifests for three years	80.5% (n = 87)	78.2% (n = 55)	Mailed	Yes
Employee training in waste handling	97.0% (n = 100)	81.3% (n = 64)	Mailed	Yes
Employee training in emergency procedures	91.7% (n = 96)	81.3% (n = 64)	Mailed	Yes
Quantities calculated to determine size	39.1% (n = 87)	15.9% (n = 63)	Mailed	Yes
Storage in labeled containers	96.8% (n = 93)	58.3% (n = 60)	Mailed	Yes
Storage in dedicated area	95.7% (n = 93)	82.0% (n = 61)	Mailed	Yes
Stored indoors or covered	95.7% (n = 93)	73.8% (n = 61)	Mailed	Yes
Storage area with impervious floor	78.9% (n = 90)	67.2% (n = 61)	Mailed	Yes
Storage area spill containment	70.4% (n = 81)	45.9% (n = 61)	Mailed	Yes
Shipped with properly licensed transporter	94.1% (n = 85)	84.0% (n = 50)	Mailed	Yes

Table 6-14. Oil and Fuel Compliance for Marinas

Performance Characteristic	Mailed Percentage	On-Site Visit Percentage	Better Performance	
All oil/fuel tanks less than 660 gallons	71.8% (n = 103)	71.7% (n = 53)	Mailed	No
Above ground capacity less than 1,320 gallons	76.6% (n = 94)	80.8% (n = 52)	On-Site	No
Below ground capacity less than 42,000 gallons	96.6% (n = 89)	100.0% (n = 52)	On-Site	Yes
SPCC on site	57.1% (n = 77)	5.9% (n = 51)	Mailed	Yes
SPCC signed by PE	27.0% (n = 63)	4.0% (n = 50)	Mailed	Yes
SPCC posted in plain view at storage locations	31.4% (n = 70)	2.0% (n = 51)	Mailed	Yes
Have above ground storage and secondary containment	77.6% (n = 58)	71.4% (n = 21)	Mailed	No
Have above ground storage and leak detection	43.6% (n = 55)	47.6% (n = 21)	On-Site	No
Have below ground storage and secondary containment	82.9% (n = 35)	95.0% (n = 40)	On-Site	Yes
Have below ground storage and leak detection	89.7% (n = 39)	95.0% (n = 40)	On-Site	No
Spill prevention procedures for receiving oil from supplier	77.0% (n = 61)	76.4% (n = 55)	Mailed	No
Spill prevention procedures for transferring within facility	70.2% (n = 47)	47.4% (n = 57)	Mailed	Yes
Spill prevention procedures for waste oil disposal	82.9% (n = 82)	67.9% (n = 56)	Mailed	Yes
Spill prevention procedures for fuel dispensing	89.4% (n = 66)	96.1% (n = 51)	On-Site	Yes
Spill containment equipment	92.5% (n = 67)	84.9% (n = 53)	Mailed	Yes

Table 6-15. Hazardous Materials Compliance for Marinas

Performance Characteristic	Mailed Percentage	On-Site Visit Percentage	Better Performance	
On-site storage calculation	62.9% (n = 62)	23.4% (n = 64)	Mailed	Yes
MSDS on file	90.7% (n = 97)	64.1% (n = 64)	Mailed	Yes
MSDS used in training	89.0% (n = 91)	52.9% (n = 68)	Mailed	Yes

Table 6-16. Storm Water Compliance for Marinas

Performance Characteristic	Mailed Percentage	On-Site Visit Percentage	Better Performance	Statistically Significant Difference
Required to have permit	27.8% (n = 72)	75.7% (n = 70)	NA [a]	Yes
Has required NPDES Permit	100.0% (n = 19)	2.4% (n = 41)	Mailed	Yes
Ablative paints: prevent material release to water	53.0% (n = 83)	60.7% (n = 61)	On-Site	No
Ablative paints: prevent material release to land	46.4% (n = 84)	57.4% (n = 61)	On-Site	Yes
Blasting, paint preparation and painting: prevent release to water	89.0% (n = 82)	84.3% (n = 51)	Mailed	No
Blasting, paint preparation and painting: prevent release to land	82.9% (n = 82)	84.3% (n = 51)	On-Site	No
Blasting, paint preparation and painting: prevent release to protect employees	100.0% (n = 79)	90.6% (n = 53)	Mailed	Yes
Engine fluids release contained	100.0% (n = 100)	96.6% (n = 58)	Mailed	Yes
Fuels, solvents, and paints stored away from drains	98.1% (n = 104)	95.1% (n = 61)	Mailed	No
Fuels, solvents, and paints plainly labeled	98.0% (n = 98)	91.8% (n = 61)	Mailed	Yes

[a] This is not purely a compliance question, but reflects the percentage that require a NPDES permit. Thus, "better performance" does not apply in this situation. A significantly smaller percentage of mailed survey respondents indicated that they did not require a NPDES permit compared to the site visit facilities. Thus, this is a concern since the site visits are being used to validate the mailed survey responses.

**Table 6-17. Summary of Compliance Question Results for Marinas—
Comparison of Mailed Surveys and On-Site Observation Compliance Rates**

Category	Number (Total Equals 38)	Percentage of Compliance Questions
Questions where mailed surveys participants had a higher compliance rate	29	76.3%
Questions where mailed survey participants had a <i>significantly</i> higher compliance rate	23	60.5%
Questions where mailed survey participants had a higher compliance rate of <i>five percentage points or more</i>	22	57.9%

Summary

The two surveys (metal finishing and marinas) had different goals and therefore yielded different results in regards to the reliability of mailed survey data. The metal finishing sector surveys asked questions on awareness and general trends in environmental performance; for this sector the mailed responses are not significantly different than the on-site observations. The marinas surveys, however, asked specific behavioral and compliance questions; the two survey modes for this sector yielded very different results.

The two surveys also asked for environmental performance data in different ways. The metal finishing survey asked for information that EPA then used to infer environmental performance (e.g., “How long is the hazardous waste maintained on site?”). The marina survey asked facilities to respond more directly as to if they were in compliance (e.g., “Are all hazardous wastes shipped with a properly licensed transporter?”). Because the questions on the marina survey required recipients to directly self-report compliance status, it is likely that the mailed marina survey would yield different results than the mailed metal finishing survey (i.e., marinas would not want to make themselves “look bad”).

E. Lessons Learned

How to Improve Response Rates

Clearly define your target population.

It is extremely important to clearly define your target population at the beginning of the project. Are you looking only for sites performing certain operations or processes? Are you looking only for a specific media of release? For example, metal finishers are typically characterized as performing plating and chemical conversion coating operations. However, this study also looked at facilities generating RCRA wastes and air emissions from other operations. Therefore, facilities that didn't perform typical metal finishing operations (e.g., painting facilities conducting surface preparation operations such as solvent degreasing) may not have considered the survey as applicable to them.

Keep your target population and mailing list data sources as local as possible.

The geographic span of your target population can impact the source of mailing data that you use. If you are targeting only a small subset of facilities, such as New England Marinas, then data can be obtained at a local level. However, if you are sending out a national, or multi-regional survey, you are less likely to have the resources to contact each locality for information. In these cases, a national database is generally required, which likely will not provide as much detail as can be obtained from local data sources. For example, for the metal finishing study EPA considered focusing on two cities, and using data from publicly owned treatment works (POTWs) and wastewater permit databases to identify survey recipients. In these cases, EPA would have been fairly certain that these facilities were in scope, since the POTWs and permit databases have data on the types of operations performed. However, by focusing on EPA Regions 1 and 5, the use of POTWs became unmanageable, so EPA used a national database (the Harris Directory) that included very general information on products and SIC code, but not specific information on the types of process performed. This likely increased the nonresponse and out-of-scope rates.

For voluntary compliance assistance programs, evaluate if the industry is receptive to EPA.

As discussed at the beginning of the section, EPA conducted the surveys for sectors in two stages of compliance assistance activity: 1) a sector for which EPA is beginning a compliance assistance effort (marinas); and 2) a sector for which EPA has conducted several compliance assistance activities (metal finishing). For both sectors, EPA and the regions coordinated with industry groups during project development. For the marina sector, EPA Region 1 was in the initial stages of a compliance assistance program, and had conducted outreach to the marinas. There were no other ongoing regulatory activities, and the marinas were receptive to the compliance assistance programs. For the metal finishers, however, EPA was in the midst of a contentious regulatory development program. Therefore, while EPA/OC worked with the trade associations to foster cooperation, some metal finishers may not have participated because of previous negative experiences with EPA.

Work with the trade associations and state compliance assistance personnel.

For both sectors, EPA worked closely with the trade associations and the state regulatory personnel in the study regions to both develop the survey instruments and foster cooperation with the industry. We believe that this communication helped improve the response rates in both sectors, and helped overcome some of the hurdles discussed above with the metal finishing sector.

Make sure the timing is right.

Because of an upcoming EPA Region 1 compliance assistance workshop, one portion of the marina survey mailout had to occur between the Christmas and New Year holidays when many businesses shut down. Very few surveys were returned as a result of this initial mailing. The entire metal finishing site visit effort was scheduled near the end of the calendar year when many businesses are at their busiest. Several facilities told the contractor that they were too busy for a site visit, but would be available at the start of the new year. These factors may have reduced the response rates. Timing your efforts to avoid major holidays or the height of a business season can help to increase the response rates of your project.

How to Choose the Right Data Gathering Technique*What type of data do you need?*

When choosing the data gathering technique give careful thought to the type of data you wish to gather. For the metal finishing survey, EPA focused on awareness of environmental requirements and general environmental performance, but did not ask for specific compliance information. As a result, the mailed survey results closely matched the on-site observations. However, for the marina survey, EPA asked more specific behavioral and compliance questions, and the results from the mailed survey did not match the results from the on-site observations. Consequently, questions pertaining to general facility information, regulatory awareness, general trends in environmental performance or customer satisfaction can reliably be obtained from a mailed survey. Whereas, if you are interested in detailed compliance or behavioral information, on-site observations are the more appropriate data gathering technique. Other data gathering techniques are also available, such as telephone or e-mail surveys, which are discussed in Section III.

Site visits provide a great deal of additional anecdotal information that can have enormous value to those running the programs. Visited metal finishing facilities were very willing to discuss their thoughts on current compliance assistance activities. Through one-on-one conversations, they often became comfortable enough to openly discuss their compliance assistance preferences, make suggestions, and even ask questions regarding areas of confusion. Also, as discussed previously, by visiting the facility the site-visit personnel were able to clarify areas of the survey instrument where the facility personnel may have been confused. Very few returned mailed surveys provided any information that was not specifically requested. When choosing the type of data gathering technique to use, you will need to decide if anecdotal information will be helpful to your effort.

Finally, during site visits, site-visit personnel were able to provide compliance assistance tools to the facility personnel. Many sites were very appreciative of the additional guidance and tools provided by the site-visit personnel. Site visits also provide the opportunity to establish a working relationship with the site. Heavily regulated industries, such as the metal finishing industry, may have a biased perception of voluntarily working with EPA. A personal, non-enforcement based visit can help to reshape those ideas and foster future cooperation.

What are your resources?

As shown in the results section above, the resources required vary greatly by the type of data that you need. Mailed surveys can be used when collecting general facility information, regulatory awareness, general trends in environmental performance or customer satisfaction type information. Site visits, while more costly, can provide detailed compliance information, additional anecdotal information, and can help to improve industry relations. Resources can dictate both the data gathering technique and the scope of the project. If resources are limited but detailed compliance information is required, a site visit program of reduced scope (i.e., non-statistical) may be the best solution.

Figure 6-1. Metal Finishing Performance Evaluation Survey

Section 1: Wastewater Treatment

1. Does your site generate chromium-bearing wastewater?
GYes **GNo**

2. Does your site generate cyanide-bearing wastewater?
GYes **GNo**

3. Does your site generate metal-bearing wastewater?
GYes **GNo**

4. Which types of wastewater treatment are used on site (check all that apply):

None		G
Equalization/Neutralization	G	
Chromium Reduction		G
Cyanide Destruction		G
Oil Skimming/Chemical Emulsion Breaking		G
Chemical Precipitation and Sedimentation		G
Complex/Chelated Metals Treatment	G	
Sand/Multimedia Filtration	G	
Membrane Filtration (Microfiltration, Ultrafiltration)		G
Ion Exchange		G
Sludge Dewatering	G	
Other: _____		

5. Do you have a wastewater discharge permit?
GYes **GNo**

6. Which activities listed below do you perform for your wastewater discharge?

Monitoring (e.g., flow, pH, ORP)		G
Sampling of wastewater for specific pollutant parameters	G	
Reporting		G

7. Have you received any compliance assistance to reduce the amount of wastewater your facility produces or to reduce the amount of metals in your wastewater?
GYes **GNo**

8. If yes, from whom? And was it helpful? _____

9. Are there any wastewater regulations that you feel you would like additional assistance with? If so, what are they? _____

10. What sources do you use for information on wastewater environmental regulations and pollution prevention? (Check all that apply)

Internet	G	Trade/professional organizations	G
Industry/professional journals	G	Environmental consultants	G
Colleagues/supervisor	G	Government environmental agencies	G

Section 2: Hazardous Waste Handling

11. Does your facility generate hazardous waste (e.g. F006 wastewater treatment sludge)?

GYes GNo

If answer to Question 11 is no, skip to Question 26.

12. Is hazardous waste generated in excess of 100 kg/month?

GYes GNo

13. How long is the hazardous waste maintained on-site?

_____ days

14. Are you aware of the hazardous waste labeling requirements for drums?

GYes GNo

15. Do they apply to your drums?

GYes GNo

16. If not, why not? _____

17. Are you aware of the hazardous waste manifest records requirements?

GYes GNo

18. If yes, do these hazardous waste manifest records apply to your facility?

GYes GNo

19. Do you perform inspections of your hazardous waste storage area?

GYes GNo

20. If yes, how often are inspections performed: _____

21. Do you have a hazardous waste contingency plan?

GYes GNo

22. Have you received any compliance assistance to help you with the identification and/or handling of hazardous waste treatment sludge?

GYes GNo

23. If yes, from whom? And was it helpful? _____

24. Are there any hazardous waste handling regulations that you feel you would like additional assistance with? If so, what are they? _____

25. What sources do you use for hazardous waste information on environmental regulations and pollution prevention? (Check all that apply)

- | | | | |
|--------------------------------|----------|-----------------------------------|----------|
| Internet | G | Trade/professional organizations | G |
| Industry/professional journals | G | Environmental consultants | G |
| Colleagues/supervisor | G | Government environmental agencies | G |

Section 3: Air Emissions

26. Does your site perform chromium electroplating or chromic acid anodizing operations?

GYes **GNo**

If answer to Question 26 is 'NO', skip to Question 34.

27. How many chromium electroplating or chromic acid anodizing tanks do you operate (excluding rinses)?

_____ tanks

28. Are you aware of emissions control requirements for your tanks?

GYes **GNo**

29. If yes, how many of your tanks use some type of emission control?

_____ tanks

30. Do you use and monitor any of the following technologies on your chromium electroplating or chromic acid anodizing solutions (check all that apply):

		Use	Monitor
Composite mesh-pad system	G	G	
Fiber-bed mist eliminator		G	G
Packed-bed scrubber		G	G
Chemical fume suppressant in plating baths		G	G
Other: _____			

31. Do you have a copy of the EPA Guidebook on How to Comply with the Chromium Electroplating and Anodizing NESHAP?

GYes **GNo**

32. Have you ever received any assistance to help you comply with EPA's Chromium MACT standards?

GYes **GNo**

33. If yes, from whom? And was it helpful? _____

34. Do you own or operate a solvent cleaning machine? **GYes** **GNo**

If answer to Question 34 is 'NO', skip to Question 44.

35. Does the solvent cleaning machine use a solvent containing 5 % or more by weight of any one or combination of the following halogenated solvents (check all solvents that are used):

Methylene Chloride	G	1,1,1-trichloroethane	G
Perchloroethylene	G	Carbon tetrachloride	G
Trichloroethylene	G	Chloroform	G

36. Which of the following solvent cleaning machines are used?

Batch Cold Machine	G	In-line Cold Machine	G
Batch Vapor Machine	G	In-line Vapor Machine	G

37. Are you aware of the EPA emissions standards for solvent cleaning machines?

GYes **GNo**

38. If yes, do these standards apply to your machines?
GYes **GNo** (Why not? _____)
39. If these standards apply, what method does your site use to comply with emission standards for solvent cleaning machines?
- Equipment compliance option **G**
 Overall emission limit **G**
 Not certain **G**
40. Do you have a copy of the EPA Guidance Document for Halogenated Solvent Cleaner NESHAP?
GYes **GNo**
41. Have you used the EPA's Halogenated Solvent Cleaner Rule Assistance Website?
GYes **GNo**
42. Have you ever received any assistance to help you comply with EPA's Solvent MACT standards?
GYes **GNo**
43. If yes, from whom? And was it helpful? _____
44. Are there any air regulations that you feel you would like additional assistance with? If so, what are they? _____
45. What sources do you use for information on air environmental regulations and pollution prevention? (Check all that apply)
- | | | | |
|--------------------------------|----------|-----------------------------------|----------|
| Internet | G | Trade/professional organizations | G |
| Industry/professional journals | G | Environmental consultants | G |
| Colleagues/supervisor | G | Government environmental agencies | G |

Section 4: Toxics Release Inventory (TRI) Reporting

46. Are you aware of the EPA Form R (TRI) for Toxic Chemical Release Reporting?
GYes **GNo**
47. Is your facility exempt from TRI reporting?
GYes **GNo** If yes, why? _____
48. Have you ever received any assistance to help you complete your TRI reporting?
GYes **GNo**
49. If yes, from whom? And was it helpful? _____
50. Are there any TRI reporting requirements that you feel you would like additional assistance with? If so, what are they? _____
51. What sources do you use for information on TRI Reporting? (Check all that apply)
- | | | | |
|--------------------------------|----------|-----------------------------------|----------|
| Internet | G | Trade/professional organizations | G |
| Industry/professional journals | G | Environmental consultants | G |
| Colleagues/supervisor | G | Government environmental agencies | G |

Section 5: General

52. Do you have internet access?

GYes **GNo**

53. Have you used the National Metal Finishing Resource Center internet site (nmfrc.org)?

GYes **GNo**

If answer to Question 53 is no, skip to question 57.

54. If yes, how do you rate the usefulness of the assistance provided?

GVery Good **GGood** **GFair** **GPoor**

55. What type of information have you obtained from the NMFRC (check all that apply)?

Compliance Information **G**
Environmental Regulations **G**
Process information **G**

56. What types of actions have you taken as a result of information you obtained on NMFRC (check all that apply)?

Changed process **G** Modified reporting procedures **G**
Implemented pollution prevention practices **G** None **G**
Implemented pollution control **G** Other: _____
Obtained permit **G**

57. Would you find the following types of compliance assistance helpful?

	Yes	No
Internet based training/guidance		G G
Workshops/Training		G G
CD ROM		G G
Onsite Technical Assistance Visits		G G
Inspections		G G
Mailed Fact Sheets/Guidance	G	G

58. Are you aware of the National Metal Finishing Strategic Goals Program?

GYes **GNo**

59. What is the number of full-time equivalent employees working at your site? _____

Figure 6-2. Marina Checklist
US Environmental Protection Agency, New England

For All Questions, Please Use:

Y: Yes

N: No

DK: Don't Know

NA: Not Applicable

Facility Type

1. a. How many employees work at the facility during peak boating recreational boating season?
Full-time ____ Part-time ____
- b. How many boats (capacity) are moored at the facility? ____ Docked? ____ Stored on land? ____
2. a. Are maintenance or repair operations performed at the facility? ____
- b. If yes, are these primarily customer do-it-yourself activities? ____

Hazardous Waste

3. a. Are manifests documenting hazardous waste shipments kept, going back at least 3 years? _____
- b. Do employees receive training in: Proper handling of wastes? ____ Emergency procedures? _____
- c. Are quantities of hazardous waste generated by the marina calculated each month, to determine what size generator the marina is? ____
- d. Are all hazardous wastes stored: In labeled containers? ____ In a dedicated storage area? ____
Indoors or covered? ____ In an area with an impervious floor? ____
With storage area spill containment? ____
- e. Are all hazardous wastes shipped with a properly licensed transporter? ____

Oil and Fuel

4. a. Is oil (including motor fuel) stored above ground in any single tank with over 660 gallons capacity? ____ Above
ground in total aggregate capacity of over 1320 gallons? ____
Below ground in total aggregate capacity of 42000 gallons? ____
- b. Is a Spill Prevention, Control and Countermeasure plan (SPCC):
On site? ____ Signed by a registered professional engineer? ____
Posted in plain view at oil storage locations? ____
5. a. Does **above** ground oil storage (including piping system) have: Secondary containment? ____
Leak detection? ____
- b. Does **below** ground oil storage (including piping system) have: Secondary containment? ____
Leak detection? ____
- c. Are spill prevention procedures in place for:
Receiving oil from a supplier? ____ Transferring oil within the facility? ____
Waste oil disposal? ____
6. a. Does the facility have spill prevention procedures in place for fuel dispensing? ____
Overfill alarm? ____ Automatic shutoff? ____ Fuel collars to capture splash/drips? ____
Employee monitoring of fueling? ____ Other? _____
- b. Is equipment available and procedures in place to contain a spill at the dispenser location? ____
- c. How would you evaluate the effectiveness of the methods you are using to prevent releases?
Excellent ____ Good ____ Fair ____ Poor ____

Hazardous Materials

7. Has the amount of each hazardous material stored onsite been calculated (including motor fuel in above-ground systems of greater than 10,000 pounds capacity) to determine if reporting to the Local Emergency Planning Committee is necessary?

8. Are Material Safety Data Sheets (MSDS) for all hazardous chemicals kept on file? _____
Used for training all employees handling hazardous chemicals? _____

Storm Water

9. a. Is a National Pollutant Discharge Elimination System (NPDES) Storm Water Permit required of this facility? _____
b. Does the marina have a NPDES Storm Water Permit? _____
10. a. When pressure washing boats coated with ablative paints, are any efforts undertaken to prevent removed material from releasing to water? ___ from contaminating land? ___
b. Are blasting, other paint preparation and painting activities contained or controlled to prevent abrasives, paint chips, and overspray from being released to the water? ___ to land? ___ to protect employees? ___
c. Are all engine fluids promptly transferred from parts, drip pans, used filters and other containers to closed receptacles for disposal or recycling? _____
d. Are fuels, solvents and paints stored in a protected, secure location, away from drains? _____
Plainly labeled? _____
11. Has the marina made structural changes to minimize surface water runoff? ___ Berming? ___
Vegetation? ___ Riprap? ___ Drains? ___ Placement of filters in drains? ___
Other? _____

Other

12. Has the facility switched to alternative materials or products to reduce toxicity or other hazards to Health, safety or the environment? ___ Safer paint stripping? ___ Safer painting? ___
Safer MSD odor chemicals? ___ Dust collection? ___ Phosphate free cleaners? ___ Biodegradable cleaners? ___
Safer antifreeze? ___ Other? _____
13. Does the facility have a sewage pumpout system? ___ Onshore ___ Boat ___ How many gallons (approx) do you pump out per week? _____ Do you use Clean Vessels Act funds? _____

APPENDICES

APPENDIX A

**PERFORMANCE PROFILE FOR
EPA's ENFORCEMENT AND COMPLIANCE ASSURANCE PROGRAM**

Impact on Environmental or Human Health Problems

Measured through annual evaluation studies of selected EPA objectives.

Effects on Behavior of Regulated Populations

Levels of Compliance in Regulated Populations

- Set 1. Rates of noncompliance for --
- a) fully-inspected populations
 - b) self-reported compliance information
 - c) populations targeted for special initiatives
 - d) priority industry sectors

Environmental or Human Health Improvements by Regulated Entities

- Set 2. Improvements resulting from EPA enforcement action
- Set 3. Improvements resulting from compliance assistance tools and initiatives
- Set 4. Improvements resulting from integrated initiatives
- Set 5. Self-policing efforts by using compliance incentive policies

Responses of Significant Violators

- Set 6. Average number of days for significant violators to return to compliance or enter enforceable plans or agreements
- Set 7. Percentage of significant violators with new or recurrent significant violations within two years of receiving previous enforcement action

Enforcement and Compliance Assurance Activities

Monitoring Compliance

- Set 8. Number of inspections, record reviews, responses to citizen complaints, and investigations conducted

Enforcing the Law

- Set 9. Number of notices issued, civil and criminal actions initiated and concluded, and self-policing settlements concluded

Providing Assistance and Information

- Set 10. Facilities/entities reached through --
- a) compliance assistance tools and initiatives
 - b) distribution of compliance information

Building Capacity

- Set 11. Capacity building efforts provided to state, local, or tribal programs

APPENDIX B

**OFFICE OF COMPLIANCE
GUIDANCE
ON THE NEED FOR
INFORMATION COLLECTION REQUESTS
(ICRS)**

**FOR THE EVALUATION OF
COMPLIANCE ASSISTANCE ACTIVITIES**

SEPTEMBER 1997

**Produced by the Office of Compliance Regional Compliance Assistance Work Group
Lynn Vendinello, Work Group Chair**

Federal agencies are generally required, by the Paperwork Reduction Act (PRA), to receive Office of Management and Budget approval prior to collecting substantially similar information from ten or more non-Federal respondents. A “collection of information” means the obtaining or soliciting of information by an agency by means of identical questions, or identical reporting or record keeping requirements, whether such collection of information is mandatory, voluntary, or required to obtain a benefit. This includes any requirement or requests to obtain, maintain, retain, report or publicly disclose information. 5 CFR § 1320.3(c)

There are exceptions to this rule and depending on your particular situation, your compliance assistance evaluation task may or may not fall within an exception. This guidance will help determine whether or not an Information Collection Request (ICR) is necessary for your task. You may also contact Hale Hawbecker in the Cross-Cutting Issue Division of the Office of General Counsel at 202-260-4555 to assist you with individual questions.

What is the Paperwork Reduction Act?

The PRA is a law (PL 104-13) originally enacted by Congress in 1980, reauthorized in 1986 and revised in 1995, that essentially attempts to minimize the Federal paperwork burden on the public. Section 3501 of the law clearly states the eleven purposes of this Act.

“ §3501. Purpose

The purposes of this chapter are to-

1. Minimize the Federal paperwork burden for individuals, small businesses, State and local governments, and other persons resulting from the collection of information by or for the Federal Government;
2. Ensure the greatest possible public benefit from and maximize the utility of information created, maintained, used, shared and disseminated by or for the Federal Government;
3. Coordinate, integrate, and to the extent practicable and appropriate, make uniform Federal information resources management policies and practices as a means to improve the productivity, efficiency and effectiveness of Government programs, including the reduction of information collection burdens on the public and the improvement of service delivery to the public;
4. Improve the quality and use of Federal information to strengthen decision making, accountability, and openness in Government and society;
5. Minimize the cost to the Federal Government of the creation, collection, maintenance, use, dissemination, and disposition of information.
6. Strengthen the partnership between the Federal Government and State, local, and tribal governments by minimizing the burden and maximizing the utility of information created, collected, maintained, used, disseminated, and retained by of for the Federal Government;
7. Provide for the dissemination of public information on a timely basis on equitable terms, and in a manner that promotes the utility of the information to the public and makes effective use of information technology;

8. Ensure that the creation, collection, maintenance, use, dissemination, and disposition of information by or for the Federal government is consistent with applicable laws, including laws relating to--(a) privacy and confidentiality,
9. Ensure the integrity, quality and utility of the Federal statistical system;
10. Ensure that information technology is acquired, used, and managed to improve performance of agency missions, including the reduction of information collection burden on the public; and
11. Improve the responsibility and accountability of OMB and all other Federal agencies to Congress and to the public for implementing the information collection review process, information resources management, and related policies and guidelines established under this chapter.”

What is an ICR?

An Information Collection Request (ICR) is a document submitted by federal agencies to OMB in order to obtain approval of an information collection and/or a reporting and record keeping requirement that falls under the purview of the PRA. The ICR must receive OMB approval prior to the initiation of the information collection.

The term “collection of information” according to the Paperwork Reduction Act of 1995 (PL 104-12(S.244)) means: “(A) the obtaining, causing to be obtained, soliciting, or requiring the disclosure to third parties or the public, of facts or opinions by or for an agency, regardless of form or format, calling for either-

“(i) answers to identical questions posed to, or identical reporting or record keeping requirements imposed on, ten or more persons, other than agencies, instrumentalities, or employees of the United States; or

“(ii) answers to questions posed to agencies, instrumentalities, or employees of the United States which are to be used for general statistical purposes.”

For guidance on how to complete an ICR, you can visit the “EPA’s Information Collection Request (ICR) Homepage” on the EPA Intranet site at <<http://intranet.epa.gov/icrintra>> or contact Monica Lewis in the Office of Policy at 202 260-1522.

When is an ICR Needed?

An ICR is generally required for any activity involving the collection of identical information from ten (10) or more non Federal respondents in any twelve month period. ICRs may be approved for up to a three-year period and can be extended through subsequent approval requests. An approved ICR is required as long as the activity continues.

Examples of information collection activities that commonly require an ICR:

- Information requirements in a rule (e.g. reporting, record keeping, waiver provisions).
- Other information collection activities (e.g. studies, surveys, application forms, audits, standardized data collection activities).

Certain activities are not subject to the Act. For example:

- An ICR is not required when the information is collected during the conduct of a criminal or civil enforcement action.
- An ICR is not needed when the collection falls under one of the categories of items that OMB has concluded do not generally meet the definition of “information” contained in 5 CFR §1320.3(h).

Many of the compliance assistance activities that the Office of Compliance is currently undertaking as well as those of the compliance assistance programs in the regions fall under one of the categories. To assist in the determination of the need for an ICR, the following examples of scenarios that do and do not require ICRs.

ICR APPLICABILITY SCENARIOS

Category A: Web -Sites:

Scenario One: I am establishing a web site for my regional compliance assistance program (or for a compliance assistance center) and would like to establish a “comments” button or “feedback” feature.

Response: Generally, no ICR would be required for this activity. According to OMB, “an undifferentiated ‘suggestion box’ format--such as one requesting ‘ideas, comments, suggestions, or anything else you would like to tell us,’ or one asking, ‘if you experience any technical problems with our site, or have any suggestions for improving it, please let us know’--are not considered to be ‘identical questions’⁴⁶. Such general solicitations of comments from the public do not require OMB clearance. However, should the agency request specific information from web site users, OMB approval would be required as explained in Scenario Two.

Scenario Two: I would like to put an on-line survey on my web site to determine what features of the web site are most useful and to ask for suggestions for improving the web site.

Response: The fact that your survey is on-line does not affect the decision as to whether or not the survey requires OMB clearance. What will affect whether or not the survey requires an ICR is the nature of your questions. According to OMB guidance, if your questions are non-identical then you will not need OMB clearance. Identical questions ask each respondent to supply the same level of information on the same subject. For example, they often supply a specific set of answers for the user to select from. Non-identical questions are non-specific and allow the responder to apply “facts or opinions” of their own choosing without any direction from the government. In addition, if your survey is primarily for the purposes of assessing customer satisfaction with your web site, you may want to consider using an existing ICR Clearance (see attached pamphlet). If your survey attempts to get at behavioral changes and/or compliance improvements, then you may need a separate clearance again depending on the nature of your questions. In general, if you feel that your questions are non-identical you may want to ensure a certain degree of brevity with respect to your survey so that your survey does not appear to follow a plan of inquiry.

⁴⁶“The Paperwork Reduction Act of 1995: Implementing Guidance”, February 3, 1997, pg. 17

Scenario Three: I would like to ask the users of my web site to identify themselves by name or by category (e.g. auto service repair shop, car dealer, consultant).

Response: No ICR would be required for this activity. According to 5 CFR 1320.3 (h)(1), this category of an inquiry is not deemed to constitute an information collection and therefore would not require clearance. The Paperwork Reduction Act states that “Affidavits, oaths, affirmations, certifications, receipts, changes of address, consents, or acknowledgment,” do not constitute information. Merely, asking users to identify themselves by name or category is a request for an “acknowledgment” not generally subject to the PRA.

Category B: Workshops/Seminars/Training Sessions

Scenario Four: I am planning to hold a compliance assistance workshop for air permits. This workshop is open to anyone who would like to attend (with limits on total numbers able to physically attend). After the workshop is over, I would like to hand out a voluntary questionnaire that asks the attendees questions such as: a) Has this workshop provided information that will help you improve your ability to comply with environmental regulations?

Response: No ICR would be required for this activity. Exemption #8 of the Paperwork Reduction Act that states that, “facts or opinions submitted in connection with public hearings or meetings”⁴⁷ would apply to this scenario. To provide for more certainty of this exemptions application, it would be best to provide a Federal Register notice making it clear that the workshop is open to all interested members of the public. A second-best option would be to adopt an open-door policy with respect to the workshop so that no one would be excluded (except for obvious space limitations) from attending. In addition, you could also conduct an on-the-spot evaluation of the workshop, since category #8 would most likely apply to that activity, as well. You could also send a follow-up questionnaire within a short time period following the seminar (e.g. one week).

Scenario Four A: My compliance assistance program has funding for four seminars this year. We would like to determine the topics that would be of the greatest interest to our clients, so we would like to mail out a voluntary questionnaire that lists potential seminar topics.

Response: An ICR would probably not be required for this activity. Category #8 of the PRA would apply to this scenario as well. OMB guidance explains that, “included in this category are questions which ask the proposed participants to identify themselves and the topic(s) about which they desire to speak.”⁴⁸ Your request for topics to be discussed is similar to asking for a request to speak on a particular topic. Further, the requested items are “in connection with” the public workshop and category #8 appears to apply to such inquiries.

Scenario Four B: After the completion of a workshop, we would like to send a follow-up survey out which asks questions about behavioral changes that resulted from attendance at the workshop.

Response: In this scenario you would probably need OMB clearance, especially if there was a significant time delay before the survey was mailed out because the information collected would no longer

⁴⁷5 CFR 1320.3 (h) (8)

⁴⁸“The Paperwork Reduction Act of 1995: Implementing Guidance”, February 3, 1997, pg 26

pertain directly to the public meeting that was held. If you were looking for behavioral changes that facilities later adopted after the workshop had educated them about environmental requirements, for example, this information request is not directly related to evaluating the immediate impact of the workshop (e.g., their satisfaction with the workshop and their improved awareness/understanding of requirements).

Scenario Five: I will be holding a printing compliance training workshop that will be made generally available to printers. I would like to administer a “test” before and after the training to determine if understanding of environmental requirements changes as a result of the training.

Response: No ICR would be required for this activity. Category #7, “Examinations designed to test the aptitude, abilities or knowledge of the persons tested and the collection of information for identification or classification in connection with such examinations,”⁴⁹ would apply to this scenario. The nature of your test should be with respect to their knowledge of the subject matter at hand. If you wish to use the test to collect socioeconomic information about the respondents, an ICR will probably be required.

Scenario Six: My office has given a grant to a state to develop a compliance guide that integrates federal and state rules for metal finishers in my state. One of the criteria for awarding the grant was that the grantee have a component for program evaluation. The grantee plans to include a comment card in the compliance guide that would get mailed back to the state office.

Response: According to OMB guidance, “In general, collections of information conducted by recipients of Federal grants do not require OMB approval. On the other hand, an agency is the sponsor of a collection of information...if the grant recipient is: 1) collecting information at our specific request; and/or 2) the terms and conditions of the grant require that the we specifically approve the collection of information or the collection procedures.”⁵⁰ One can ask for a program evaluation component of a grant proposal and/or measures of success; however, we can not ask that a particular survey method be used without getting an ICR approved.

If, however, the award is a cooperative agreement, then the agency is considered a sponsor of the information and all of the PRA restrictions on information collection would apply.

Category C: Mailed or Phoned Surveys

Scenario Seven: An EPA employee or contractor would like to know how many states have a small business policy and plans to call them to ask for a copy of their policy, if they have one.

Response: No ICR would be necessary to conduct this activity. Category #2 of the Paperwork Reduction Act states that the request for “samples of products or any other physical objects”⁵¹ does not constitute a collection of information. According to OMB, this category “includes requests for information that is already available in a form suitable for distribution and is provided in that form to all

⁴⁹5 CFR 1320.3(h)(7)

⁵⁰IBID, pg 14

⁵¹5 CFR 1320.3(h)(2)

requesters. (This request is a collection of information if the information has to be compiled or if it is not provided to any person who requests it).⁵²

Scenario Eight: An EPA employee or contractor would like to follow up with those states that have sent us a copy of their small business policy to ask specific questions about their individual state policies.

Response: No ICR would be required for this activity. Since you will be asking each of the states questions that pertain only to their specific policy and not identical questions of each state, Category #6 of the PRA would apply. Category #6 of the PRA states that “a request for facts or opinions addressed to a single person”⁵³ does not constitute a request for information. However, if EPA asked the same questions following a plan to more than nine states, the PRA would apply.

Scenario Nine: My EPA program would like to ask states to voluntarily answer a survey that asks them to quantify the benefits of their compliance assistance program.

Response: It is important to understand that the PRA applies not only to industry and individuals but also to requests for information from states and local governments. Further, the fact that the survey is voluntary does not mean that the PRA does not apply. In this case OMB clearance would be required because you are asking identical questions and are directing them to specific entities.

⁵²“The Paperwork Reduction Act of 1995: Implementing Guidance,” draft, February 3, 1995 pg. 24

⁵³5 CFR 1320.3(h)(6)

APPENDIX C

**How to Obtain Clearance
for
Regional Compliance Assistance Evaluation Surveys
Under the Generic ICR 1860.01
OMB Control # 2020.0015**

EPA

**U.S. Environmental Protection Agency
Office of Compliance
Office of Enforcement and Compliance Assurance**

Table of Contents

Steps to Developing the Clearance Package in a Glance 1

The Memos 2-3

Request for Approval of Information Collection Activity 4-5

The Survey 6-7

Frequently Asked Questions

 Question 1: When do I need to use the generic compliance assistance ICR? 8

 Question 2: Is an ICR required for all surveys? 8

 Question 3: How do I receive approval for my survey, if it meets the conditions outlined above? 8

 Question 4: How Long Will The Process Take? 8

 Question 5: What are the new requirements in the latest ICR? 9

 Question 6: Can I Send a List of Survey Questions That I Plan to Use over and over Again at
 Workshops, Training Seminars, or On-site observations to OMB for Approval? 10

 Question 7: What Do I Have to Do to My Survey to Show That it Has Been Approved by
 OMB? 10

 Question 8: Is There a Standard Cover Memo for the Clearance Package and, How Do I
 Calculate the Necessary Burden Estimates? 10

Steps to Developing the Clearance Package in a Glance

- Write a brief memo from YOUR program office to Monica Lewis, OECA Desk Officer

From: Your Name

Office Name

To: Monica Lewis, OECA Desk Officer

Collection Strategies Division

Office of Environmental Information

In the memo include:

Survey title

Originating office, point of contact with phone number

State whether your ICR is under the 10 day review (no behavior change questions) or the 30 day review (survey has behavior change questions)

Number of expected survey respondents

Cost and Burden estimate

- **Write a brief memo from Monica Lewis, OECA Desk Officer to Bryon Allen, OIRA Desk Officer for OMB.**

From: Monica Lewis, OECA Desk Officer

Collection Strategies Division

Office of Environmental Information

To: Bryon Allen, OIRA Desk Officer

Office of Management and Budget

In the memo include:

Survey title

Originating office, point of contact with phone number

State whether your ICR is under the 10 day review (no behavior change questions) or the 30 day review (survey has behavior change questions)

Number of expected survey respondents

Cost and Burden estimate

- **Develop the Request for Approval of Information Collection Activity**

I. Background

II. Survey Purpose and Description

III. Survey Methodology and Use of Results

IV. Respondents Burden

V. Agency Burden

- **Include a Copy of the Survey** - remember to include the reporting burden statement at the end

**EXAMPLES AND INSTRUCTIONS
FOR ALL THE DOCUMENTS FOLLOW**

EXAMPLE

Memo One

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION I
1 CONGRESS STREET, SUITE 1100
BOSTON, MASSACHUSETTS 02114-2023

MEMORANDUM

DATE: July 17, 2000

SUBJ: Request for OMB Approval of a Survey for Viewers of the Video RCRA Compliance for Metal Finishers.

FROM: Linda Darveau
New England Environmental Assistance Team

TO: Monica Lewis, OECA Desk Officer
Collection Strategies Division
Office of Environmental Information

I am the EPA project officer on EPA Grant # X991962 to the Northeast Waste Management Official Association, which is producing a 22-minute videotape for metal finishers entitled RCRA Compliance for Metal Finishers. I would like to issue a mail-back survey to assess viewers' opinions of the video. Since this survey contains behavioral change questions, it will necessitate the maximum 30-day OMB review under the generic ICR.

I plan to offer the video free of charge by advertisement to approximately 1500 metal finishers and other business and government contacts in New England. I estimate that we will distribute 500 videos. Each video will be accompanied by a postage-paid return survey to assess how informative and useful the viewers find the video. Based on a 20% response rate, I expect to receive 100 responses. This, in combination with the brevity of the survey, means that the burden to both EPA and the public will be low.

If you have any questions or concerns, please call me at 617-xxx-xxxx.

Attachments (2)

CC: Lynn Vendinello
Rochele Kadish

EXAMPLE**Memo Two****MEMORANDUM**

SUBJECT: Review of Compliance Assistance Regional Program Evaluation Survey,
ICR No. 1860.01 (OMB 2020-0015)

FROM: Monica Lewis, OECA Desk Officer (mailcode 2822)
Regulatory Information Division
Office of Policy

TO: Bryon Allen, OIRA Desk Officer
Office of Management and Budget

As a condition of OMB approval for the generic Regional Compliance Assistance Program Evaluations ICR, No. 1860.01, Expiration date January, 2005, OMB Control No. 2020-0015, EPA agreed to submit each specific questionnaire covered by this clearance to OMB for review. Attached for your review and approval is a Region 3 survey for evaluating the results of a mailing. The purpose of this survey is to assess the effectiveness of the mailings in helping facilities to comply, informing them about EPA's policy for waiving or reducing penalties for disclosed violations, and determining possible areas for EPA improvement. Since this survey contains behavioral change questions, it will necessitate the maximum 30-day OMB review under the generic ICR.

Note: EPA will not have access to names or addresses of facilities surveyed and will not use the information for inspection or enforcement purposes. There will be no way to link participant responses to specific facilities or individuals.

Your comments and suggestions would be much appreciated. For comments or questions on the survey, please contact Janet Viniski at 215-xxx-xxxx.

If you have any questions about the ICR or this survey submission, please contact me at 202-566-1678

Thank you for your cooperation in this matter.

Attachments

EXAMPLE

Request for Approval of Information Collection Activity

I. Background:

Under EPA grant X991962 to NEWMOA, EPA Region I is producing a 22-minute videotape for metal finishers entitled RCRA Compliance for Metal Finishers in the Northeast. The video is designed to provide basic up-to-date information on RCRA Compliance to small facilities whose personnel have difficulty sparing the time to attend workshops and read documents. The video includes information on both regulatory requirements and pollution prevention.. The video will convey this information in an appealing, easy-to-digest manner by showing narrated views of real processes, focusing on the most common violations found by both state and EPA inspectors. The video will be offered free of charge by advertisement to 1300 metal finishers and other business and government contacts in New England. I estimate that we will distribute 500 videos. Each video will be accompanied by a postage-paid return survey to assess how informative and useful the viewers find the video.

II. Survey Purpose and Description:

I have drafted a postage-paid mail-back survey to accompany each video that is distributed. The purpose of the survey is to assess how informative and useful the video is to members of the intended audience in helping them understand and make plans to act upon pollution prevention opportunities. EPA will use this information to improve future videos for this and other sectors and to choose appropriate outreach methods (videos versus documents versus workshops, etc).

III. Survey Methodology and Use of Results

Each video requested and mailed out will be accompanied by the survey. Of the 1300 metal finishers and other business and government contacts in New England to whom we will advertise the video free-of-charge, I estimate that 500 will request a video. Based on a 20% response rate I expect to receive 100 responses. Each survey will be printed as a postage-paid self-mailer so that respondents do not have to provide envelopes or postage. EPA Region I will be the mailing address to which the surveys are returned, and Linda Darveau, the EPA contact for this project, will compile the results and provide them to interested parties– EPA and state pollution prevention assistance programs, Metal Finishing trade associations, and others.

IV. Respondents' Burden

Number of survey recipients: 500

Number of respondents: 100

Minutes per response: 10

Cost per hour: \$33

(Based on the Bureau of Labor Statistics "Employer Costs for Employment Compensation," the estimate used in the EPA / OECA generic ICR.)

Total burden hours: 16 hours

Total burden dollar cost: \$549.00

V. Agency Burden

Number of survey recipients: 500

Number of respondents: 100

Minutes spent compiling paper responses, per response: 10

Burden hour subtotal for paper responses: 16

Additional hours to summarize and communicate all results: 15

Burden hour total: 31

Cost per hour: \$41.25

(Based on a 2002 GS 13/01 salary of \$64,542 or \$31.73/hr; with 30% overhead the hourly rate is \$41.25. This is the estimate used in the EPA / OECA generic ICR.)

Total burden dollar cost: \$1278.75

EXAMPLE Survey

1. Overall, did you find the video worthwhile?

Yes Somewhat No

2. How helpful are the following portions of the video?

Role of Inspectors:	Very	Somewhat	Not
Training Requirements:	Very	Somewhat	Not
Waste Inspection Requirements:	Very	Somewhat	Not
Hazardous Waste Determination:	Very	Somewhat	Not
Hazardous Waste Storage:	Very	Somewhat	Not
Pollution Prevention:	Very	Somewhat	Not

3. Do you intend to make any changes at your facility as a result of viewing this video?

Yes No

If yes, what do you intend to do? _____

4. Do you intend to refer to the video in the future or pass it on to a colleague?

Yes No

5. Which of the following ways of receiving information on environmental issues work best for you?

(Check any.)

- Manuals (long, comprehensive)
- Fact sheets (short, single-topic)
- Newsletter or journal articles
- CD Roms
- Internet Web Pages
- Workshops
- Videos
- Other (please specify)

6. What type of job do you have?

- Owner of a metal finishing facility.
- Environmental Manager of a metal finishing facility.
- Waste water Treatment Operator.
- Consultant.
- Government Assistance Provider.
- Other (please specify): _____

7. If you plan to test or adopt a new practice or technology at your facility as a result of this video, and may be willing to talk to us about your experience, please write your name and telephone number below. We ask because we want to know if our non-regulatory assistance efforts result in real improvements. Thank you!

8. Other comments or suggestions: _____

Thank you for your input!

**Linda Darveau
The New England Environmental Assistance Team
EPA Region I**

Public reporting burden for this collection of information is estimated to average ten minutes per response, including time for reviewing instructions, gathering information, and completing and reviewing the information. Send comments on the Agency's need for this information, the accuracy of the provided burden estimate, and any suggestions for reducing the burden, including the use of automated collection techniques, to the Director, Office of Environmental Information, Collection Strategies Division, United States Environmental Protection Agency (Mail Code 2822), 1200 Pennsylvania Ave, NW Washington, D.C. 20460; and to the Office of Information & Regulatory Affairs, Office of Management & Budget, 725 17th Street, NW, Washington D.C. 20503, Attention: Desk Officer for EPA. Include the EPA ICR 1860.01 and the OMB control number 2020-0015 in any correspondence. Do not send your completed survey to this address. Approval expires January 2005.

Frequently Asked Questions

Question 1: When do I need to use the generic compliance assistance ICR?

The generic clearance is to be used for **voluntary collections** of program evaluation information when EPA wants to **survey more than nine** recipients of a compliance assistance activity. This generic ICR can be used for the bulk of the planned program evaluation/outcome measurement work planned in the regions and headquarters, where your goal is to determine the effectiveness of your compliance assistance activities on the audience that receives the compliance assistance (e.g., participants at a workshop). Programs are encouraged to use good survey design, methodology, and implementation in all of their program evaluation work.

Question 2: Is an ICR required for all surveys? An ICR is not required for surveys or pre/post tests that are given during a workshop that has been well publicized - open to the public.

Question 3: How do I receive approval for my survey, if it meets the conditions outlined above?

Prior to initiating the survey, sponsoring regional programs must seek final survey approval from OMB. To obtain approval, the sponsoring regional program must **submit a clearance package** (a full example is on pages 2-4) consisting of a memorandum and a copy of the survey instrument through the Collection Strategies Division (CSD) in the Office of Environmental Information.

CSD staff will review each submission to ensure that it meets the requirements of the Paperwork Reduction Act and the generic approval, and may reject any proposed survey that does not meet those requirements.

Question 4: How Long Will The Process Take?

Approximately 45 days at the most and 14 days at the least - this is a combination of EPA and OMB time.

EPA process: Send your package to Lynn Vendinello and Rochele Kadish. They will do a quick review (1-2 days) and send it on to Monica Lewis who will submit the survey and attached materials to OMB.

For those surveys with behavioral change questions, approximately 15 days after the survey has been sent to OMB, you should ask Lynn Vendinello in OC to send an e-mail to OMB to remind them of their 30 day deadline. In addition, you can also state in your cover note to OMB that if you do not hear back from OMB within the required time, you will assume that they have no comments on the survey and that it is approved.

OMB process: will review surveys that **evaluate satisfaction** with or improved understanding that results from compliance assistance **within 10 working days**. OMB will review surveys that **evaluate behavioral responses** to compliance assistance (i.e., asking one or more of the five questions allowed to evaluate behavioral change) **within 30 working days**.

Question 5: What are the new requirements in the latest ICR? The only new restrictions for EPA is the limitation of behavior change questions. You can only ask the following questions:

1. What regulatory actions do you intend to take (did take) as a result of the <compliance assistance activity>?
Check all that apply:
 - a. File a notification
 - b. Obtain a permit or certification
 - c. Provide data to EPA, state, or local regulator
 - d. Submit reports to EPA, state, or local regulator
 - e. No regulatory actions will be taken
 - f. Other specify _____

2. What process changes at your facility do you intend to make (did make) as a result of the <compliance assistance activity>?
Check all that apply:
 - a. Change the storage or handling of a waste or emission
 - b. Change a process or practice
 - c. Purchase new process equipment
 - d. Implement material or waste recycling system
 - e. Install pollution control equipment (e.g., scrubbers, control technique)
 - f. Install a waste treatment system
 - g. No process changes will be taken
 - h. Other specify _____

3. What management changes at your facility do you intend to take (did take) as a result of the <compliance assistance activity>?
Check all that apply:
 - a. Conduct a self audit
 - b. Establish an environmental management system
 - c. Institute training or other communication to improve awareness and/or practices
 - d. No management action will be taken
 - e. Other specify _____

4. Who (if anyone) have you contacted (will contact) for further assistance as a result of the <compliance assistance activity>?
Check all that apply
 - a. A vendor
 - b. A state or local regulatory agency
 - c. A non-regulatory local source for additional compliance assistance
 - d. No contacts will be made
 - e. Other specify _____

5. Of the activities that you have implemented as a result of the <compliance assistance activity> what have been the resultant pollutant reductions and cost savings, if applicable?

Activity	Pollution/Reduction	Cost Savings
----------	---------------------	--------------

Question 6: Can I Send a List of Survey Questions That I Plan to Use over and over Again at Workshops, Training Seminars, or On-site observations to OMB for Approval?

Yes. If you are going to be repeating your compliance assistance activities over the course of the year(s) you can send a comprehensive list of survey questions for pre-approval. The list of questions would be broader than the list for any one survey. For example, you may know that you will be conducting MACT training seminars for wood finishers, drycleaners and paint coaters next year. You also know that you want to follow-up on the seminars by asking those that attended what changes they made at the facility as a result of the training. You would develop your list of questions that applied across all three sectors plus those specific to each sector for OMB approval. When you delivered the surveys you would only use those questions that were appropriate to the audience being assisted. In calculating burden, however, you will need to base it on the overall number of respondents to all of the surveys (see BURDEN calculation example below).

Question 7: What Do I Have to Do to My Survey to Show That it Has Been Approved by OMB?

The OMB Control Number and expiration date must appear on the front page of an OMB-approved form or survey, or on the first screen viewed by the respondent for an on-line application. The rest of the burden statement must be included somewhere on the form, questionnaire, or other data collection tool, or in the instructions or cover letter for such collection.

The following information must appear on the first page of the survey:

OMB Control No. 2020-0015

Approval expires January 2005

Public reporting for this collection of information is estimated to average XX minutes per response, including time for reviewing instructions, gather information, and completing and reviewing the information. Send comments on the Agency's need for this information, the accuracy of the provided burden estimate, and any suggestions reducing the burden, including the use of automated collection techniques to the Director, Office of Environmental Information, 1200 Pennsylvania Ave., NW Washington DC 20460 and to the Office of Information and Regulatory Affairs, Office of Management and Budget, 725 17th St. NW Washington DC 20503. Attention: Desk Officer for EPA. Include the EPA ICR 1860.01 and the OMB control number 2020-0015 in any correspondence. Do not send your completed surveys to this address.

Question 8: Is There a Standard Cover Memo for the Clearance Package and, How Do I Calculate the Necessary Burden Estimates?

Yes, a sample memorandum is on page 2. In the memo you need to address five areas:

- 1) **Background:** Briefly describe the compliance assistance activity being undertaken;
- 2) **Survey Purpose and Description:** Briefly describe the parameters that you are attempting to measure through the survey and attach the actual survey (or list of survey questions);

3) **Survey Methodology and Use of Results:** Explain how you plan to conduct the survey (e.g., mail or phone survey, onsite visit) and how you will use the results;

4) **Respondents Burden:** Calculate total respondents' burden hours and total respondents' cost. Follow the approach below and see the example in Attachment 1.

a) **Number of Respondents:** How many respondents do you anticipate responding to the survey.

b) **Minutes per response:** How long will it take to complete the survey. In the ICR, our estimates per respondent were as follows: phone surveys-10 minutes; mailed/faxed back surveys-10 minutes; pre/post tests-10 minutes; onsite revisits-120 minutes; and online surveys- 5 minutes.

c) **Cost per hour:** In the ICR, we estimated the cost per respondent at \$32.68 plus a 110 percent overhead for a total hourly rate of **\$68.63**. This rate applies to both the regulated community and state and local technical assistance staff. This figure is based on data from the Bureau of Labor Statistics (May 2001, Employer Costs for Employment Compensation) for the typical types of small businesses that we serve through our compliance assistance program. You should use this figure in calculating respondents' cost.

Total Burden Hours = (# of Respondents X Minutes per response)/60

Total Respondents Cost = \$68.63 X Total Burden Hours

5) **Agency Burden:** This refers to the time that it will take you to review the responses and conduct your analysis. You will need to supply a Total Agency Burden estimate, which is calculated by multiplying EPA staff time (in hours) by cost per hour. In the ICR, we estimated cost per hour as **\$41.25**. This rate is based on the 2002 GS-13-01 annual salary of \$64,542 or \$ 31.73/hr, plus a 30 percent overhead. This figure should be used in your calculation of Agency Burden.

If you supply OMB with a list of survey questions that you will use throughout your regional compliance assistance program, you will still need to perform the above calculations. However, the estimate should reflect the total anticipated burden for all of the surveys that use questions from the list.

APPENDIX D

MENU OF SAMPLE SURVEY QUESTIONS BY OUTCOME MEASURE

This appendix provides a menu of sample survey questions you can tailor for your specific evaluation efforts. Use them as a guide to get you started. Add questions, as appropriate, and delete questions that are unrelated to your evaluation effort. Modify the questions to fit the sector you targeted as well as for the type of compliance assistance activity. If your survey is conducted under the generic ICR, questions indicating behavioral and environmental changes must be limited to the five listed in Section I, parts B and C below. Additional survey questions can be found at :

<http://www.epa.gov/compliance/planning/results/surveys/index.html>

Section I of this document focuses on outcome measurement, which is central to OECA's commitments under the Government Performance and Results Act. Section II lists supplemental questions for onsite visits and Web sites. Section III of this document includes sample questions to assess the background of the respondent and customer satisfaction.

SECTION I: OUTCOME MEASUREMENT

Measuring the outcomes of compliance assistance—changes in understanding, behavioral change, and environmental and human health impacts—is a central component of OECA's performance measurement strategy. Questions related to the key measures that OECA plans to track are listed in bold text. Table 1 lists the outcome measures.

Table 1. Outcome Measures

Outcome Measure	
Understanding	Number of facilities/respondents who say they better understand how to comply with environmental regulations
	Change in level of understanding (as measured through pre and post tests)
Behavioral Change	

Outcome Measure	Specific Measures
Environmental and Human Health Improvements	Number of facilities that reduced emissions or other pollutants
	Number of facilities indicating a safer environment for workers
	Quantified environmental improvements (e.g., amount of emissions or pollutants reduced)

A. Outcome Measurement Category: Understanding

1. Would you say that you are more knowledgeable about environmental requirements and opportunities as a result of this compliance assistance?

- Yes
 No
 N/A

2. As a result of the assistance you received, how has your understanding of the environmental regulations that apply to your business improved?

- A great deal. I feel that I understand what is required.
 Somewhat. I am still a bit confused about the regulations.
 Not at all.
 N/A

Comments: _____

3. What would have helped you to understand the environmental regulations more fully?

- More clearly written regulations.
 Better written guidance materials.
 A more knowledgeable staff person.
 A training class or workshop.
 More time to read the materials.
 Other: _____

4. What did you learn that will be most useful to you?

- How to apply for a permit.
 Information on new equipment or techniques to use to lower emissions.
 How to implement an environmental management system.
 The name of a contact in another regulatory department.
 Information on how similar companies have reduced emissions or improved compliance.
 Other: _____

B. Outcome Measurement Category: Behavioral Change

If your survey is conducted under the generic ICR, questions indicating behavior change are limited to the four below:

1. What regulatory actions do you intend to take (did take) as a result of the <compliance assistance activity>?

Check all that apply:

- File a notification.
- Obtain a permit or certification.
- Provide data to EPA, state, or local regulator.
- Submit reports to EPA, state, or local regulator.
- No regulatory actions will be taken.
- Other specify: _____

2. What process changes at your facility do you intend to make (did make) as a result of the <compliance assistance activity>?

Check all that apply:

- Change the storage or handling of a waste or emission.
- Change a process or practice.
- Purchase new process equipment.
- Implement material or waste recycling system.
- Install pollution control equipment (e.g., scrubbers, control technique).
- Install a waste treatment system.
- No process changes will be taken.
- Other specify: _____

3. What management changes at your facility do you intend to take (did take) as a result of the <compliance assistance activity>?

Check all that apply:

- Conduct a self audit.
- Institute an environmental management policy, system or procedure.
- Institute training or other communication to improve awareness and/or practices.
- No management action will be taken .
- Other specify: _____

4. Who (if anyone) have you contacted (will contact) someone for further assistance as a result of the <compliance assistance activity>?

Check all that apply

- A vendor.
- A state or local regulatory agency.
- A non-regulatory local source for additional compliance assistance.
- No contacts will be made.
- Other specify: _____

C. Outcome Measurement Category: Environmental and Human Health Improvements

- 1. Of the activities that you have implemented as a result of the <compliance assistance activity> what have been the resultant pollutant reductions and cost savings, if applicable?.

Activity	Pollution/Reduction	Cost Savings
_____	_____	_____
_____	_____	_____
_____	_____	_____

SECTION II: SUPPLEMENTAL QUESTIONS FOR ONSITE WORKSHOPS AND WEB SITES

A. Supplemental Questions for Workshops

- 1. Was the material presented clearly and in a logical sequence?

- Yes
- No

- 2. How would you rate the handouts and materials?

- Excellent
- Very Good
- Good
- Fair
- Poor

- 3. Would you like more seminars like this one made available to you?

- Yes
- No

If yes, please list topics:

- 4. What was the most useful part(s) of the workshop?

- 5. What was the least useful part(s) of the workshop?

- 6. What topic(s) would you have liked to have spent more time on:

7. Would you be willing to spend more than 2 hours at a workshop in order to cover more topics?
- Yes
- No

8. Do you have any suggestions to help us reach more people like yourself to have them attend these seminars?

9. Was the location of this meeting convenient for you?

- Yes
- No
- Somewhat

Comments: _____

10. What other locations would you recommend for future seminars?

11. Was the time of the seminar convenient for you?

- Yes
- No
- Somewhat

12. What other times might be more convenient?

- Morning (8:00 a.m. - 11:00 a.m.)
- Lunch (12 noon - 2:00 p.m.)
- Afternoon (2:00 p.m. - 5:00 p.m.)
- Evening (7:00 p.m. - 9:00 p.m.)

13. Did the speakers show knowledge of the subject?

- Yes
- No

14. Was the technical level right for you?

- Yes
- No

15. Were the questions handled appropriately?

Yes

No

16. What did you learn that will be the most helpful to you?

17. On a scale of 1-10 (10 being the highest) how would you rate:

_____ The workshop

_____ The presenters

18. Would you be interested in having a followup onsite compliance assessment?

Yes

No

B. Supplemental Question for Web Sites

1. How did you learn about this Web site?

Search engine (please specify): _____

Link from another Web site (please specify): _____

From an EPA document

Referral from a colleague

Other (please specify): _____

SECTION III: SUPPLEMENTAL BACKGROUND AND CUSTOMER SATISFACTION QUESTIONS**A. Supplemental Background Questions**

1. What type of organization do you work for?

- Regulated facility or business
Industry sector: _____
- Consulting company or law firm
- Government
- Trade association
- Nonprofit organization
- School or university

2. How did you become aware of this [insert name of compliance assistance activity]?

- Referral from another government agency, official, or hotline
- Referral from another business
- Trade association
- EPA letter or mailing
- EPA workshop, seminar, or conference
- Web site
- EPA publication or newsletter

3. In the past, have you used any other compliance assistance tools provided by EPA, such as (check all that apply):

- Hotline
- Fact sheets
- Guidance documents
- Web site
- Onsite visits
- Workshops, seminars, or conferences

APPENDIX E

Examples of Letters Used by OECA to Gather Data by the Dillman Method

Example Metal Finishing Prenotification Letter for Survey Recipients

Dear Metal Finisher:

Over the past several years, EPA has conducted several compliance assistance efforts targeted to metal finishing facilities to improve environmental quality and compliance rates. EPA has recently begun to measure the results of these efforts and to identify whether additional compliance assistance is needed.

To support these efforts, EPA will be conducting a study of metal finishers in EPA Regions 1 and 5⁴ to identify areas where additional compliance assistance is needed, to identify the most useful types of compliance assistance, and to develop a “snapshot” of current performance towards key Federal environmental regulations. EPA will use both mailed surveys and site visits to conduct this study in order to compare the efficiency and effectiveness of these two data gathering techniques. ***This study will be a voluntary blind study. This means that EPA will NOT know the identity of the facilities participating in the study. This study is NOT intended to provide information on specific facilities, but instead is intended to develop aggregate statistics on a subset of the industry to evaluate compliance assistance needs. There will NOT be any enforcement follow-up activities related to this study.***

EPA has contracted with Eastern Research Group, Inc. (ERG), to conduct the study. ERG has randomly selected 100 metal finishers in EPA Regions 1 and 5 for site visits, and has randomly selected another 200 facilities to receive the mailed survey. This letter is a notification that your facility has been randomly selected to receive a mailed survey. Your participation is voluntary; however, by participating you will provide EPA with information on how EPA can improve compliance assistance delivered to metal finishers, and you will help EPA measure the effectiveness of its compliance assistance projects. As stated above, this is a ***blind*** study; EPA will not know which facilities participate in the study. All surveys will be identified by code; facility names or addresses will not be included on the surveys. ERG randomly selected the facilities to participate in this study from an industry population identified through three data sources: the Harris Directory, EPA Regions 1 and 5, and Publicly Owned Treatment Works.

Within the next one to two weeks, you will receive a mailed survey developed by EPA to evaluate awareness of and performance towards the following key Federal environmental regulations: the solvent degreasing NESHAP standard, the chrome MACT standard, Clean Water Act requirements, RCRA requirements, and EPCRA reporting requirements. The survey also includes several questions to assess what types of compliance assistance have been used by the facility, and what types of compliance assistance facilities find most useful.

ERG has established a helpline to support any facilities needing additional information or guidance on completing the checklist. The helpline number is 1-866-867-4637. In addition, EPA will provide a self-addressed stamped envelope for returning the survey. To help maintain the confidentiality of the responses, please do not add a return address to this envelope.

To maintain our project schedule, EPA will be requesting that you complete and return the survey within two weeks of receipt. The results of this study will be published on EPA’s Web site and will also be made available to the Association of Electroplaters and Surface Finishers (AESF). If you have any questions or concerns, please call the helpline.

EPA recognizes that you have many demands on your time. This study will help EPA develop and deliver better focused compliance assistance tools that we hope will save you time and money in the future. We thank you in advance for participating in this study.

Sincerely,

Linda Darveau
Environmental Specialist
US EPA New England
(617) xxx-xxxx

⁵⁴EPA Region 1 comprises Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont, while EPA Region 5 comprises Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin.

Example Metal Finishing Prenotification Letter for Site Visits Recipients

Dear Metal Finisher:

Over the past several years, EPA has conducted several compliance assistance efforts targeted to metal finishing facilities to improve environmental quality and compliance rates. EPA has recently begun to measure the results of these efforts and to identify whether additional compliance assistance is needed.

To support these efforts, EPA will be conducting a study of metal finishers in EPA Regions 1 and 5⁵⁵ to identify areas where additional compliance assistance is needed, to identify the most useful types of compliance assistance, and to develop a “snapshot” of current performance towards key Federal environmental regulations. EPA will use both mailed surveys and site visits to conduct this study in order to compare the efficiency and effectiveness of these two data gathering techniques. ***This study will be a voluntary blind study. This means that EPA will NOT know the identity of the facilities participating in the study. This study is NOT intended to provide information on specific facilities, but instead is intended to develop aggregate statistics on a subset of the industry to evaluate compliance assistance needs. There will NOT be any enforcement follow-up activities related to this study.***

EPA has contracted with Eastern Research Group, Inc. (ERG), to conduct the study. ERG has randomly selected 100 metal finishers in EPA Regions 1 and 5 for site visits, and has randomly selected another 200 facilities to receive the mailed survey. This letter is a notification that your facility has been randomly selected for a site visit. Your participation is voluntary; however, by participating you will provide EPA with information on how EPA can improve compliance assistance delivered to metal finishers, and you will help EPA measure the effectiveness of its compliance assistance projects.

As stated above, this is a ***blind*** study; EPA will not know which facilities are visited. All site visit information will be identified by code; facility names or addresses will not be included in any reports. ERG randomly selected the facilities to participate in this study from an industry population identified through three data sources: the Harris Directory, EPA Regions 1 and 5, and Publicly Owned Treatment Works.

Within the next week, you will receive a phone call from ERG to determine the best available date and time to schedule the site visit. During the 2-4 hour site visit, an ERG employee will complete a copy of the survey developed by EPA to evaluate awareness of and performance towards the following key Federal environmental regulations: the solvent degreasing NESHAP standard, the chrome MACT standard, Clean Water Act requirements, RCRA requirements, and EPCRA reporting requirements. The survey also includes several questions to assess what types of compliance assistance have been used by the facility, and what types of compliance assistance facilities find most useful. The results of this study will be published on EPA’s Web site and will also be made available to the Association of Electroplaters and Surface Finishers (AESF).

EPA recognizes that you have many demands on your time. This study will help EPA develop and deliver better focused compliance assistance tools that we hope will save you time and money in the future. We thank you in advance for participating in this study.

Sincerely,

Uylaine Barringer
Assistant Program Manager
RCRA Compliance

⁵⁵EPA Region 1 comprises Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont, while EPA Region 5 comprises Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin.

Example Metal Finishing Cover Letter for Survey Recipients

Dear Metal Finisher:

Within the past two weeks, you should have received a letter notifying you of an EPA study of metal finishers to identify areas where additional assistance is needed related to compliance with environmental regulations, to identify the most useful types of compliance assistance, and to develop a “snapshot” of current performance towards key Federal environmental regulations. ***This study is a voluntary blind study. This means that EPA will NOT know the identity of the facilities participating in the study. This study is NOT intended to provide information on specific facilities, but instead is intended to develop aggregate statistics on a subset of the industry to evaluate compliance assistance needs. There will NOT be any enforcement follow-up activities related to this study.*** Your facility has been randomly selected to participate in this study. Your participation is voluntary; however, by participating you will provide EPA with information on how EPA can improve compliance assistance delivered to metal finishers, and you will help EPA measure the effectiveness of its compliance assistance projects.

HOW DO YOU PARTICIPATE?

To support EPA in this effort, we ask that you:

- 1) Complete the attached survey. The survey is designed to evaluate awareness of and performance towards key Federal environmental regulations, and to assess what types of compliance assistance you have found useful.
- 2) Call the survey toll-free helpline number at 1-866-867-4637 if you have any questions on how to complete the survey.
- 3) Place the survey in the enclosed self-addressed stamped envelope and return it by two weeks after receipt.

EPA recognizes that you have many demands on your time. This study will help EPA develop and deliver better focused compliance assistance tools that we hope will save you time and money in the future. We thank you in advance for your time in participating in this study.

Sincerely,

Linda Darveau
Environmental Specialist
US EPA New England
(617) xxx-xxxx

Example Metal Finishing Reminder Postcard for Survey Recipients

The following text appeared on a 5.5" × 8" postcard made from blue cardstock.

Metal Finishing Performance Survey Reminder

Dear Metal Finisher,

Please remember to complete the EPA Metal Finishing Performance Evaluation Survey and return it in the enclosed, self-addressed stamped envelope by December 14th. If you have any questions regarding survey completion, please contact the toll-free helpline at 1-866-867-4637.

If you have already returned your completed survey, thank you for your participation.

Sincerely,

Linda Darveau
US EPA New England

Example Metal Finishing Re-mail Cover Letter for Survey Recipients

Dear Metal Finisher:

In late October, you should have received a survey regarding an EPA study of metal finishers to identify areas where additional assistance is needed related to compliance with environmental regulations, to identify the most useful types of compliance assistance, and to develop a “snapshot” of current performance towards key Federal environmental regulations. ***This study is a voluntary blind study. This means that EPA will NOT know the identity of the facilities participating in the study. This study is NOT intended to provide information on specific facilities, but instead is intended to develop aggregate statistics on a subset of the industry to evaluate compliance assistance needs. There will NOT be any enforcement follow-up activities related to this study.*** We have not yet received your survey. Your participation is voluntary; however, by participating you will provide EPA with information on how EPA can improve compliance assistance delivered to metal finishers, and you will help EPA measure the effectiveness of its compliance assistance projects.

HOW DO YOU PARTICIPATE?

To support EPA in this effort, we ask that you:

- 1) Complete the attached survey. The survey is designed to evaluate awareness of and performance towards key Federal environmental regulations, and to assess what types of compliance assistance you have found useful.
- 2) Call the survey toll-free helpline number at 1-866-867-4637 if you have any questions on how to complete the survey.
- 3) Place the survey in the enclosed self-addressed stamped envelope and return it by two weeks after receipt.

EPA recognizes that you have many demands on your time. This study will help EPA develop and deliver better focused compliance assistance tools that we hope will save you time and money in the future. We thank you in advance for your time in participating in this study.

Sincerely,

Linda Darveau
Environmental Specialist
US EPA New England
(617) xxx-xxxx



APPENDIX F

Advanced Data Analysis –Measuring Association and Making Comparisons

The descriptive statistics discussed in Section V above were only concerned with measuring characteristics of one variable at a time. In this section we discuss measure of association and how to make statistical comparisons (i.e., hypothesis testing) when conducting data analysis. We expect that each of the methods discussed in this section will be accessible to individuals with basic math skills. We describe some more advanced techniques below. The more advanced techniques may require additional support from a qualified statistician.

Measures of Association

Measures of association deal with how two or more variables relate to one another. There are two measures of association that you should be aware of: correlation coefficients and contingency tables.

Correlation coefficients measure the direction and strength of the relationship between two variables. They are most useful when both of the variables you are concerned with can take on unlimited values.⁵⁶ Due to complicated nature of the calculation, we exclude the formula for calculating correlation. Most spreadsheets, however, have a correlation function that is relatively easy to apply. The following can assist you in interpreting correlation coefficients:

1. A correlation coefficient can take on values between -1.0 and 1.0.
2. Negative values of the correlation coefficient indicate an inverse relationship between the two variables: i.e., as one variable is increasing in value, the other is decreasing in value.
3. Positive values of the correlation coefficient indicate a positive relationship between the two variables: i.e., as the value of one variable increase, the value of the other variable also increases.
4. Correlation coefficients that are close to (or equal to) zero indicate no correlation between the two variables.
5. Larger values of the correlation coefficient (in absolute terms) indicate stronger correlations. For example, a correlation of 0.65 indicates a stronger relationship than a correlation of 0.35.⁵⁷

A correlation of 1.0 or -1.0 should be interpreted with caution. It is rare to find two variables that are perfectly correlated with one another. If this occurs, you should make sure that one variable is not being derived from the other.⁵⁸ Another caution you should consider is that correlation does not imply causation.

⁵⁶ See Section V's definition of limited variables.

⁵⁷ The same is true for negative correlations, a -0.65 is a stronger correlation than a -0.35.

⁵⁸ For example, if we had information on the pounds of a chemical used by a company and the total cost of that chemical, we should find a perfect positive correlation if we use the same dollar amount for each observation to derive the cost.

That is, a positive correlation indicates that when one variable takes on a large value, so does the other variable. It does not imply that the large value of one variable caused the other variable to also be large.

When you have three or more variables that you are interested in, a “correlation matrix” is a useful means of displaying the correlation coefficients. To form such a matrix, simply write the variable names across the column headings and then (in the same order) down the row headings. Then, fill in the correlation coefficients for each pairing below the diagonal.⁵⁹ The following table provides an example of what a four-variable correlation matrix would look like.

Variables	A	B	C	D
A	1.0	-	-	-
B	0.85	1.0	-	-
C	0.65	0.75	1.0	-
D	0.01	0.02	-0.25	1.0

The above-diagonal cells can be left blank since they are redundant with the below-diagonal cells. Note that in the example, variables A, B, and C are all strongly positively correlated with one another. Variable D, on the other hand, is not correlated with A and B and only weakly and negatively correlated with C.

The second form of measuring association are contingency tables, or cross-tabulations. Contingency tables are useful when you are interested in looking at discrete or limited-value variables. With a contingency table you can look at the relationship between the values for two or more variables. There are no set ways of developing contingency tables, so we provide a few examples. The key, however, is to form a set of columns from the answers to one or more questions and form rows from the answers to one or more different questions.

Suppose you are interested in how small and large entities differed in their answers to a variety of survey questions regarding the usefulness of certain compliance assistance tools. The following table could help you look at that question, taking into account that not all entities accessed all the tools.

Tools:	Tool A		Tool B		Tool C	
Facility Size	Used Tool	Found it Useful	Used Tool	Found it Useful	Used Tool	Found it Useful
Small (305 total)	305	102	305	98	202	198
Large (151 total)	151	34	62	57	47	44
TOTAL	456	136	367	155	249	242

⁵⁹ The diagonal of the matrix is each variable correlated with itself, which by definition should equal one.

This table combines information on how many accessed each tool, the size of those that accessed each tool, and how many (of both size classes) that found each tool useful. As can be seen in the table, Tool A was not very useful but Tool C was useful to respondents. The usefulness of Tool B differed between small and large facilities with large facilities finding it more useful than small facilities.

As another example, suppose you are interested in the relationship between the answers to two survey questions: one that asked respondents to rate their satisfaction with a compliance assistance program on a scale of one to five (five = very satisfied) and another that asked respondents to identify which compliance assistance tools they had used (Tools A, B, and C, as above). A contingency table for this question might look like this:

Satisfaction With Program		Tools Accessed By Those Respondents		
Level	Total Responses	A	B	C
1	56	56	34	5
2	79	79	65	19
3	159	159	131	75
4	125	125	102	118
5	37	37	35	32
TOTAL	456	456	367	249

In this example, we can see a clear pattern: satisfaction with the program increased with use of Tool C and to a lesser degree with Tool B. No additional information on the relationship between use of Tool A and satisfaction is possible, however, since all respondents accessed Tool A.

Making Comparisons (Hypothesis Testing)

Hypothesis testing is the backbone of inferential statistics and as such is one the most important aspects of statistical analysis. In its most basic form, hypothesis testing attempts to determine if an observed value or relationship is attributable to random chance. Values or relationships that are not attributable to random chance are generally referred to as statistically significant. In this discussion we provide some basic information on when hypothesis testing should be employed and we also provide information on a few simple tests that you can use. For the most part, complicated hypothesis tests should be left to those with good statistical backgrounds.

Hypothesis tests should be employed when you want to verify that a certain value or relationship is not due to simple random chance. For example, you may want to test whether the compliance rate among respondents was significantly greater than 50 percent. You may want to determine whether small and large entities had the same rate of satisfaction with a compliance assistance program. As a final example, you may want to determine whether a calculated correlations coefficient was significantly different from zero. In other words, you will want to conduct hypothesis tests if you want to make inferences (statements) about your data (or the underlying population) with some confidence.

To simplify our discussion of hypothesis testing, we focus on four tests that should be most useful to you:

- # *One Sample Test for a Proportion*—A test to determine if a percentage is significantly different than a specific value.
- # *Two Sample Test for a Proportion*—A test to determine if a percentage calculated from one sample is significantly different than a percentage calculated from another sample.
- # *One Sample Test for a Mean*—A test to determine if an average value is significantly different than a specific value.
- # *Two Sample Test for a Mean*—A test to determine if an average calculated from one sample is significantly different than an average calculated from another sample.

As can be seen, the four tests cover one and two samples and proportions and averages. In a one sample test, we are testing to see if the value in the sample is significantly different than a specific value. In a two-sample test, we have values calculated from two different samples and we are looking a significant difference between the two values. In the two-sample test, the two samples can be sub-groups from one larger sample (e.g., larger versus small facilities). A proportion is any variable whose “mean” value is expressed as a percentage (e.g., the percentage of facilities that responded with a “yes” to certain question). Mean values are simply non-percentage values for the means of variables. We present the formulas for these tests in Table F-1.

Table F-1. Equations for Performing Simple Hypothesis Tests

	Proportions	Means
One-Sample	$z = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}}$	$z = \frac{(x_s - \mu)}{\frac{s}{\sqrt{n}}}$
Two-Sample	$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$	$z = \frac{(x_1 - x_2)}{\sqrt{\frac{s_1}{n_1} + \frac{s_2}{n_2}}}$
Definitions	<p>z—A statistical parameter used to test for significance.</p> <p>p_s—The estimated value of the proportion from the sample.</p> <p>p—The value that you want to compare (statistically) p_s to.</p> <p>n—The number of observations in the sample.</p> <p>p_1, p_2—The estimated proportion values for samples 1 and 2 respectively.</p> <p>n_1, n_2—The number of observations in samples 1 and 2 respectively.</p>	<p>z—A statistical parameter used to test for significance.</p> <p>x_s—The estimated mean value from the sample.</p> <p>μ—The value that you want to compare (statistically) x_s to.</p> <p>n—The number of observations in the sample.</p> <p>x_1, x_2—The estimated mean values for samples 1 and 2 respectively.</p> <p>n_1, n_2—The number of observations in samples 1 and 2 respectively.</p>
Significance Criteria	<p>If the absolute value of z is:</p> <ul style="list-style-type: none"> -less than 1.65, then the difference is not statistically significant. -between 1.65 and 1.96, then the difference is significant at the <i>ten</i> percent level of significance. -between 1.96 and 2.58, then the difference is significant at the <i>five</i> percent level of significance. -greater than 2.58, then the difference is significant at the <i>one</i> percent level of significance. <p>Note: The absolute value is calculated by simply turning a negative number positive or by making no adjustment to a positive number.</p>	

In the table, we define a few “significance criteria” which tells you how to translate the values you estimate into a level of significance. These criteria result in one of four levels of significance: not significant, 10 percent, five percent, and one percent. The percent levels refer to a level of confidence in the results. For example, the five percent level of significance implies that there is only a five percent chance that your hypothesis test was “wrong”. That is, there is some chance that you obtained a sample that is not representative of the true population and thus your test has given you a false answer. To interpret the results of your tests, you can use the following criteria defined in Table F-2.

Table F-2. Interpreting Hypothesis Test Results

Test	Interpretation
One Sample Test for a Proportion	A significant difference (at any level) implies that your estimated sample proportion (p_s) is significantly different than the test value of p . The <i>relative</i> magnitudes of the two values will tell you whether the sample proportion is significantly greater than or less than the test value.
Two Sample Test for a Proportion	A significant difference (at any level) implies that the estimated proportion from the first sample (p_1) is significantly different than the estimated proportion from the second sample (p_2). The <i>relative</i> magnitudes of the two values will tell you which one is significantly greater.
One Sample Test for a Mean	A significant difference (at any level) implies that your estimated sample mean (x_s) is significantly different than the test value of μ . The <i>relative</i> magnitudes of the two values will tell you whether the sample mean is significantly greater than or less than the test value.
Two Sample Test for a Mean	A significant difference (at any level) implies that the estimated mean from the first sample (x_1) is significantly different than the estimated mean from the second sample (x_2). The <i>relative</i> magnitudes of the two values will tell you which one is significantly greater.

Advanced Data Analysis—Additional Techniques to Analyze Data

In this section we introduce some more advanced techniques that can be used to look at relationship and trends in your data. We present these techniques in only summary form since properly applying them will require someone that has experience in applying these more advanced techniques. In our summaries, we attempt to provide some indication of what each of these techniques can be used for and what they might be able to provide you in terms of final results.

Analysis of Variance (ANOVA)—In an ANOVA analysis, the idea is to determine the influence that different factors have had on some variable. Usually the factors of interest are *qualitative* in nature. For example, suppose you have conducted some compliance assistance seminars to improve understanding and you asked participants to take a pre- and post-seminar test. The seminar was conducted by three different outside instructors and you conducted the seminar in two different ways. After conducting all of the seminars, you want to look at the improvement in test scores (from the pre- to the post-seminar test). There is one complication, however: you have three different instructors and two different ways in which the seminar was conducted. An ANOVA analysis can assist you in sorting through the influence of these different factors. In this case, an ANOVA will tell you which factors had the biggest influence toward increasing scores between the pre- and post-seminar test.

Regression Analysis—The purpose of a regression analysis is to look at the influence of *quantities* of different variables on one variable. Thus, while ANOVA looks the influence of different qualitative factors, a regression analysis would look at the influence of quantitative factors on a specific variable. For example, suppose you had information on the amount of chemicals used by a set of facilities, as well

as a number of other quantitative factors that you think may “explain” (i.e., influence) the variation in chemical usage (e.g., capacity of final product, etc).⁶⁰ A regression analysis would allow you to estimate the quantitative relationship between each explanatory factor and chemical usage.

Analysis of Category Data—Categorical data analysis is an extension of the contingency tables discussed above. This form of analysis is useful when the relationships you are interested in are between variables that are categorical in nature (e.g., EPA Region, states, type of process, compliance status, compliance assistance tools used, etc). For example, suppose you have information on the compliance status of several entities and the types of compliance assistance tools each has used (including none). Suppose you are interested in which tools are associated with higher levels of compliance. A categorical data analysis would provide this type of information by performing a hypothesis test to determine if compliance status is related to use of different tools. The analysis could tell you which tools are associated with higher levels of compliance.

Drawing Appropriate Conclusions

In performing statistical analysis one must be careful to draw appropriate conclusions from the results. In particular, there are three issues that you should consider when you are drawing conclusions.

First, you should be aware that statistical results that are based on a sample are subject to some degree of error. That is, since the results are not based on the whole population, you are only generating *estimates* which you then assume are representative of the population. Given this, there is still some chance that you drew an unrepresentative sample and that your results are not accurate. This can happen even if you follow all of the prescription in Section V on sampling. Thus, it is necessary to remember that your results may be wrong. You should always report the size of your sample, the size of the population that it is meant to represent, how you drew the sample or collected the data, and non-response rates for surveys.

Second, statistical analyses are subject to the principle of “garbage in, garbage out.” A sophisticated statistical analysis will not overcome bad data. For example, you may hire a statistician to perform several advanced analyses of survey data that you collected to answer a number of important questions. However, if the survey data is flawed (e.g., a non-random sample, too few observations collected, rampant nonresponse), then the resulting analyses will also be flawed. Thus, your analysis can only be as good as your data.

Finally, statistical analyses do not *prove causation*. A statistical analysis can only demonstrate tendencies or how two or more variable relate to one another. For example, suppose you are looking at the relationship between a set of compliance assistance tools and improvements in compliance status, and you find that one tool is strongly related to improved compliance. This does not *prove* that the use of the tool *caused* an improvement in compliance. The strongest statement that you can make is that the tool is strongly related to improved compliance. There may be other factors not accounted for in the analysis, or the relationship between the tool and improved compliance may more complicated than a simple cause-effect relationship. In other words, there may be other explanations for the observed relationship that the statistical analysis cannot test for or encompass. Thus, it is necessary not to place an interpretation on your results that the analytical framework cannot support.

⁶⁰ Actually, not all of the factors need to be quantitative. Regression analysis can also encompass some qualitative factors also.



APPENDIX G

Sampling From Unknown Populations (Draft from 06/13/02)

The Challenge

One of the most important components of any survey project, whether you are conducting on-site visits, a paper survey, an Internet questionnaire or a phone survey, are the people who answer your survey. For your survey to be valuable you need to reach a set of people who are representative of the overall group that interests you. In statistical terminology, the set of people you contact is your **sample** and the whole group is the **population**. You do not need to contact everyone in the population since you can obtain reliable and accurate information from a sample while substantially reducing your data collection cost. For example, let's say there are a total of 1,000 dry cleaners in your area but you're only going to survey 100. In this case, the 1,000 is the total population and 100 is the sample. Additionally, if you construct your project correctly, the 100 dry-cleaners you survey should provide accurate and reliable information on the characteristics and practices of the 1,000 dry cleaners in the population.

The difficulty arises when the population size is unknown or you have little information on who is in the population. In this situation, you need to develop a reliable and accurate population list from which to draw the sample of survey participants. This applies to both personal interview projects as well as written surveys. Without a reliable and accurate list, your results may be skewed or biased. In these situations, developing this list is one valuable outcome of the project.

Purpose of the Guidance

This guidance is intended to assist you in developing a *reliable* population list, from which to derive (1) a random sample for statistically valid measurements or (2) an informal sample for non-statistically valid measurements. A reliable list produces accurate information on as much of the intended population as possible without including entities that are not part of the population.

We begin by discussing the characteristics of populations that are generally unknown or for which reliable lists are typically not available (Section I). We then turn to potential data sources and ideas for developing population lists (Section II). In our discussion of data sources, we try to give some insight into how you develop reliable lists from multiple sources. Section III discusses how to estimate the population size from the lists you have developed. It is important to estimate your population size before you take measurements from that population. Section IV provides some simple ideas on how to draw samples from these populations to obtain reliable and useful information.

I. Characteristics of Unknown Populations

Several characteristics are common among unknown populations. These characteristics stem from the reasons *why the businesses are not found on population lists*. For sectors dominated by businesses that fit these characteristics, *definitive* population lists are rare. The characteristics that we list below are certainly not comprehensive, but should provide you with an idea of why some populations are unknown.

Sectors dominated by small businesses—Sectors with *a majority of small businesses* (e.g., “Mom and Pop” operations, high attrition rate businesses, small annual sales) tend to have unknown

populations. *These businesses have fewer resources, limiting their participation in trade organizations.* Also, in a physical sense these businesses are less noticeable since they are smaller. Finally, groups of small businesses may be less likely to form trade organizations than groups of larger businesses.

Less-formalized or informal businesses—There are a number of *sectors where some (or most) of the businesses operate in an informal manner.* For example, the home remodeling sector contains a range of business organizations from incorporated contractors to handyman services, and the auto salvage sector includes many “hobbyists.” It may be relatively easy to find lists of more formal business operations (e.g., contractors), but less easy to find lists of less formal operations. In some cases, the less formal operations may be actively seeking to avoid detection for tax purposes.

Businesses without extra resources—Businesses operating with extremely *tight margins* frequently do not pay to participate in trade associations or other types of listing activities. Sectors where this is common will thus tend to have unknown populations.

Situations where the population definition and business sectors do not match perfectly—For example, you may be interested in chromium electroplaters, a form of metal finishing. Chromium electroplaters, however, are not usually identified separately from other metal finishers. Thus, a list of metal finishers that doesn’t identify the type of electroplating operations may be too broad.

Sectors where there are language barriers—Sectors where the majority of businesses do not speak English or where English is spoken as a second language may also tend to have unknown populations. In these cases, language works as a barrier to the development of population lists.

Seasonal or economy-driven sectors—There are numerous sectors where the season or the state of the economy either determines their existence or heavily influences the number in the sector. Finding reliable lists for these populations can be challenging since businesses within the sector come and go frequently.

II. Data Sources

Table 1 lists several good data sources to develop population lists. In Table 2, we provide some comments on other sources that may *not* be as useful⁶¹. You should consider the following as you develop your population list:

Use more than one source—For unknown populations, you should use more than one source in developing your population list. *As you increase the number of sources, you are increasing the reliability of your list.* For one, coverage may differ between sources with some sources missing certain groups of the population while others may contain those groups. Using multiple sources will result in overlap; you should take note of overlapping entities.⁶² In using multiple sources, you should track the sources for each entity that you identify.

⁶¹ The purpose of Table 2 is to steer you away from the sources in that table.

⁶² Overlapping can also be seen as a measure of reliability. Specifically, if you have two sources that you feel are fairly accurate but somewhat incomplete, you can be fairly certain that the overlapping entities are part of the population.

Evaluate each source—In evaluating the usefulness and reliability of any source you need to determine *how the source developed its list*, including the reliability of the data used. You need to know if inclusion in the list required some form of payment or fee.⁶³ Finally, you need to know when the list was last updated since older lists tend to be less accurate or complete. You should make notes of all of this information to better assess the reliability of your final list.

Identify any biases that may result from the sources you use—An important aspect of developing your population list is to recognize that the resulting list may be biased. For example, you may have developed a population list for auto repair shops from a trade association list, the yellow pages, and a marketing data service. Each of these may be biased in some way. For example, the trade association may require annual dues, limiting the membership list to shops that can afford to be members, and not everyone pays to be included in the yellow pages. *Reviewing how the source list was developed will allow you to determine what biases may exist in your final list.* For example, all of your sources may be biased against inclusion of smaller facilities. Using multiple sources should reduce the bias in your final list.

Look at the total acquisition cost—A final consideration in developing your list is to *consider the trade-off between dollars and labor hours in developing population lists.* Each source will require some resources to acquire. Some sources can be obtained by simply purchasing the data from a company that provides the information (e.g., Dun and Bradstreet, a company that specializes in collecting data on public and private companies). Other sources may be free, but may require substantial labor effort (on your part) to verify the reliability and accuracy of the data. In considering any potential source you should consider what you are getting from the source (reliability and completeness) compared to what you are paying for that source (dollars and/or labor hours).

⁶³ For example, a trade association may publish a list of members, but may require a fee from the members for inclusion in the list.

Table 1. Some Potential Sources for Developing Population Lists

Source	Comments
Trade Associations	Numerous trade associations operate in the United States, many of which maintain comprehensive member lists. Although there are several directories of trade associations, the more comprehensive ones are commercial products and it may not be cost-effective to purchase one if all you need are names of a few trade associations. Many libraries have the trade association directories in print form. One common and comprehensive directory is the <i>National Trade and Professional Associations of the United States</i> (Columbia Books, published annually). Additionally, you can use the Internet to locate relevant trade associations. One consideration for trade associations is to determine if all members of the association will be part of your target group.
Phone Book	The local-area phone book is a reliable source for locating businesses. CD-ROMS that contain Yellow Page listings are available and offer flexible search options. You should determine the extent to which your targeted group fits into the categories offered in the Yellow Pages.
Federal or State Agency Lists	A number of Federal and state agencies maintain lists of companies for various reasons. OSHA, for example, maintains a database of all inspected companies (the Inspection Management Information System (IMIS)). One thing you should be aware of though is that inclusion in these lists often depends on a number of criteria. For example, inclusion on OSHA's IMIS list requires an inspection and that the company employ more than 10 people. (The OSHA IMIS database can be accessed at http://www.osha.gov/oshstats/index.html .) In addition, state and local agencies may also have regulatory interaction with the population of interest, and may have population lists. For example, state Departments of Transportation or Motor Vehicles may have lists of auto salvage yards.
Data Collection Companies	There are some data collection companies that primarily serve the financial sector. The most prominent among these companies are Dun and Bradstreet (D&B) and Standard and Poors' (S&P) Compustat. The data from both of these covers a large number of private and public companies. The companies do not, however, tend to cover smaller companies. Nevertheless, these sources may be useful. Contacting either company and discussing the issue with a sales representative will assist you in determining if they can be of assistance. (Dun and Bradstreet: 800-624-5669; Compustat: 800-523-4534).
Business Licensing Departments	All states require some form of licensing for businesses to operate. You should consider contacting these departments to determine the level of information they can provide. Some states are more willing than others to share information.

Table 1 (continued)

Source	Comments
Direct Marketing Lists	<p>Several companies offer lists of businesses that can be used in direct marketing campaigns. These companies draw their lists from several sources (e.g., Yellow Pages, directories). Some of them offer customized services. These companies typically charge a fee for lists; the fee should be weighed against the benefits of quickly obtaining reliable lists. Before using such a company, however, you should review the sources that the company uses and the likelihood that any list you purchase will match your target group. A good resource for finding these companies is the <i>DM News</i> web site (http://www.dmnews.com). The site contains a “Yellow Pages” of direct marketing companies that lists companies by the services they provide. We suggest companies in the “List Broker/Manager/Compiler” or “Mailing List Broker” categories as a starting point.</p>
Drive-bys for limited population sizes and areas	<p>One manual means of collecting data is through a “drive-by” approach. If you are interested in a limited geographic area (e.g., a specific city, a specific body of water) then collecting the population information by physically enumerating them may be possible. That is, you go out into the field and record each population member that you come across. This process, however, can become complicated as the size or the geographic area increases in size or complexity. Nevertheless, if you can efficiently cover the geographic area and can accurately identify population members, then this approach may be an effective means of developing a list.</p>
Equipment manufacturers and suppliers	<p>Most businesses require some form of equipment and supplies to operate. For example, many auto repair shops purchase parts from auto parts dealers/retailers. Some equipment manufacturers and suppliers may be willing to provide detailed customer lists. These lists may provide a basis for a population list. The key here is to find manufacturers or suppliers that provide specialized equipment (i.e., used only by the population you are interested in) <i>and</i> that provide the equipment to a large number of the population. There are some considerations here, however. Specifically, some members of the population who purchase from these sources may have asked not to have their names and information passed on to third parties. Another consideration is that not all members of the population may purchase from the same manufacturer or supplier. Nevertheless, if a manufacturer or supplier is willing to share a customer list, they may also be willing to discuss the completeness and accuracy of that list.</p>

Table 2. Some Sources That Are *Not* As Valuable For Building Population Lists

Source	Comments
Marketing-Oriented Lists	Several sources offer lists of companies providing a specific service. For example, you may find a list of “dry cleaner supply companies” that has been distributed to dry cleaners. You should determine how that list was compiled. There are cases where lists are compiled for marketing purposes. Inclusion on such lists may require payment. This would mean that these lists are biased towards those that are willing to pay to be on the list.
Better Business Bureau	The Better Business Bureau (BBB) is a valuable resource for a number of things, but not as a source for population lists. BBB data do not include sufficient detail and rely on a number of potentially biasing criteria.
Chambers of Commerce	As with the BBB, Chambers of Commerce offer valuable resources to the business community. Chambers of Commerce are not, however, good sources of population lists since membership in these organizations generally requires payment of dues as well as other requirements.
U.S. Census Bureau	The U.S. Census Bureau is an excellent source for aggregated data, but is not a source for company-specific data. The Census will not release any information that can identify a specific company. Census Bureau data is useful, however, in estimating the population size as we will discuss below.
Trade Show Participants	Companies that have participated in trade shows may seem like a good starting point, but refining the list may be more trouble than it is worth. Trade shows offer a meeting place for a specific service or product category. Even if you are able to find one that caters specifically to your target group, be aware that not all participants will be in the target group. Many participants may be offering services or products that are used by your target group. Trade show participant lists may be useful to have, but may not provide useful input into a population list.

III. Estimating the Population Size

When dealing with unknown populations, developing an estimate of the population size is an important first step. The problem, however, is that there is no standard method for estimating the population size in cases of unknown populations. This section presents ideas on how to estimate population sizes using available information. We offer three approaches to estimating population size and a method of combining the three approaches to develop a “reasonable range” for your population size.

Approach 1: Use the Number in the List You Compiled—Your first consideration should be to assess the completeness and accuracy of your final list. For this assessment, consider how each individual source developed its own list, determine which segments of your population may be missing from your final list, and identify if out-of-scope entities are in your list. *If you can say with some certainty that your final list is relatively complete and accurate, then your best estimate of your population size is the number of entities in your list.* This is rarely the case, however, with unknown populations.

Approach 2: Use Source Estimates of Coverage to Generate a Population Number—A second approach to estimating the population size is to discuss the accuracy of a source with the organization that developed the source. This is certainly more productive with some sources (e.g., trade associations) than with others (yellow pages). Organizations such as trade associations or marketing data services may have already developed estimates of their coverage. For example, a trade association may be able to tell you that they cover 75 percent of all businesses in the sector. From there you can make an estimate of population size.⁶⁴ *You should obtain an estimate of coverage from each source that can provide a reasonable estimate of coverage* (e.g., the phone company may not be able to provide such an estimate for the yellow pages). *If you can do this for more than one source, then look at the range of those estimates.* If that range seems reasonable, then the range can work as a feasible estimate of the population size.

Approach 3: Census Bureau Data—A third means of the estimating the population size is to use data collected by the U.S. Census Bureau. The U.S. Census Bureau collects information on businesses operating in the United States. *The Census data that may be most useful for regional and state-level performance measurement projects would be the County Business Patterns data,⁶⁵ which presents the numbers of business establishments for various sectors at the state, county, and metropolitan statistical area (MSA) levels.* For example, using this data for 1999, there are 3,828 “automotive repair and maintenance shops” in Massachusetts, of which 89 percent employ less than 10 employees. Sectors are defined using the North American Industrial Classification System (NAICS) which provide a great amount of detail. For example, “automotive repair and maintenance shops” are further divided into mechanical and electrical repair and maintenance and auto body repair. The mechanical and electrical grouping is then further divided into other groups. The census web site also offers detailed definitions of the NAICS, allowing you to determine which ones are relevant to your project.

⁶⁴ In this case, the estimate is to take the number in the trade association list and divide by 0.75.

⁶⁵ The Census *County Business Patterns* data can be accessed online at <http://www.census.gov/epcd/cbp/view/cbpview.html>. It is available in an online database, HTML, and PDF formats.

Approach 4: Combination of above three—Finally, you can combine all three of the above approaches and *formulate a “reasonable range.”* A range is possibly the best you can do for cases of unknown populations. A reasonable range should include the actual population size, while a single estimate risks being far away from the actual value since we have little information on the actual population size. Thus, we suggest using a range as your size estimate unless you are relatively certain you are close with a single estimate.

To formulate the reasonable range, we suggest you write down all of the estimates that you have made from the above-mentioned approaches. Then, rate each on a scale of one to five on how confident you are in that estimate, with five being the most confident. You can then use the following rules to develop a “reasonable range”:

If you have more than one “five”—Use the smallest and largest estimates with a five rating as your range.

If you have one “five”—Use your most confident value below and above the estimate rated as “five” as your range. Describe it as a range with a “most likely value” equal to the five-rated value. If the five-rated value is not within the range, use it as either the upper or lower bound of the range and describe it as “a range of x to y, but most likely closer to y” (assuming y was the five-rated value).

If you have no “fives”—Use your most confident lowest and highest values as your range.

You should try not to use any “ones” or “twos” in developing your range. For example, in the second rule above, you may be better off using the five-rated value as an upper or lower bound rather than using a “one” or “two” as the upper or lower bound. If applying this general rule, however, results in only one remaining value, then you are better off using some “ones” and “twos” than having only a single estimate.

This method is not based on theoretical statistical concepts, but is intended for someone with limited statistical background. Statisticians have developed means of estimating population sizes for unknown populations based on theoretical statistics. However, unless you have access to a qualified statistician, successful implementation of these ideas may be impossible. Also, the three approaches used for this method should not be considered the only means of estimating population size. Any other approach that provides a reasonable estimate is worthwhile exploring.

IV. Sampling Methods

This section discusses three *informal sampling methods* that can be used select members of the population for inclusion in your study: **judgement sampling, quasi-random sampling, and purposive sampling**. We call these informal methods because they do not satisfy the requirements of random sampling. This does not, however, imply that these methods provide invalid or misleading results. *Informal sampling methods can provide useful and insightful information at less cost than statistically valid methods*. The ability to extrapolate the results from informal methods to a larger population, however, is limited since you cannot ensure the representativeness of your sample. For each method, we assume that you have been able to compile a fairly reliable list and that you have made some estimate of the population size. We provide some examples of how to conduct these type of sampling exercises in Table 3. Following our discussion of the three sampling methods, we provide a short discussion of choosing a sample size.

Judgement Sampling

Judgement sampling involves choosing a sample **based on your ideas of who should be in the sample**. This method of sampling is the farthest from statistically valid random sampling since it involves a subjective choice for selecting the sample. The advantage of this method, however, is that when choosing the participants, you can ensure that your sample contains respondents who can provide interesting or detailed information. For example, you may select a facility that you know has a detailed environmental management system and has addressed issues in which you are interested. You may want to choose facilities that fit a certain profile (e.g., large facilities that perform a specific operation). The key here is that you set the criteria for inclusion based on specific reasons.

The problem with judgement sampling is that the sample you obtain will not represent the entire population. Thus, taking information from your judgement sample and extrapolating this to your population is not valid or reliable. On the other hand, by specifying who you talk to, you can maximize the amount of information you obtain. **This type of sampling method can assist you in defining issues, identifying problems or solutions, and understanding processes.**

Quasi-Random Sampling

A quasi-random sample involves choosing a **subjective criterion but using a probability-based mechanism for choosing participants**. The subjective criterion can be any number of things, but it usually has to do with convenience. For example, a quasi-random sample may be a survey taken of seminar participants that asks questions beyond the topics of the seminar. Another example of a quasi-random sample would be to query every 10th member of the population.

Quasi-random samples are not representative and may provide results that are biased in an unknown manner. Quasi-random samples, however, are a less-costly means of selecting a sample than a judgment sample, and are also useful when you have little information on who to choose in a judgement sample.

Table 3. Examples of the Three Informal Sampling Methods

Method	Examples	Best Applications
Judgement Sampling	<ul style="list-style-type: none"> • Choose “best performing” facilities. • Choose facilities that have compliance problems. • Choose facilities that have been active in other Agency programs. • Choose facilities that perform specific operations. • Choose facilities that have been helpful in the past. 	<p>Judgement sampling is best suited for situations where you need specific or specialized information. That is, you are less concerned with obtaining a representative group of participants, but more concerned with selecting a group that can tell you specific things (e.g., how a compliance assistance program improved environmental performance).</p>
Quasi-Random Sampling	<ul style="list-style-type: none"> • Choose every <i>n</i>-th (e.g., tenth, fifth) on the list. • Randomly sort your list using a random number generator (e.g., from a spreadsheet) and choose participants from the start of the sorted list until you reach your target number of participants. • Choose seminar participants as your sample. • Choose facilities that are convenient for on-site interviews. • For a mailed survey, choose some of the facilities on your list. 	<p>Quasi-random sampling is best suited for situations where you want to obtain a somewhat representative sample, but do not have the resources to develop a statistically valid approach. A quasi-random sample sacrifices some randomness (and hence representativeness) to reduce the cost of the collection. Nevertheless, some effort is made to obtain a sample that is representative, given the resource constraint.</p>
Purposive Sampling	<ul style="list-style-type: none"> • Divide the population between some relevant groups (e.g., between small and large) and choose a sample that mimics that division. For example, if 20 percent of the entire population is small, then be sure 20 percent of your sample is small). • Choose some facilities with poor compliance records and some with stellar compliance records. Divide the sample evenly between the two groups, if possible. 	<p>Purposive sampling is best suited for cases where one or more factors (e.g., size, compliance status, use of a specific compliance assistance tool) is important. In most cases where purposive sampling is used, you are making sure that you have some information on each “side” of the factor (e.g., both large and small) because you are particularly interested in exploring differences between the two “sides.”</p>

Purposive Samples

A purposive sample uses **objective selection criteria**, but a **non-probability mechanism** to choose participants. In a purposive sample, you are looking to replicate some aspect of the population, which can be very complicated (e.g., trying to replicate statistical parameters). One form of purposive sampling involves replicating population characteristics. For example, you may know that 80 percent of all facilities are “small” facilities. If you had a sample size of 20, you would choose four “large” facilities and 16 “small” ones.⁶⁶

Purposive samples are more difficult to select and require more information on the population. However, purposive samples are the least biased and are closer to statistically valid than the other two methods. They are not, however, representative since they do not satisfy the requirements of statistical validity.

Choosing A Sample Size

Given that these methods are not statistically valid, the number of participants that you choose can be based on resource constraints. Naturally, as the number of participants rises, the more reliable your results. To determine a sample size, we suggest that you determine the sampling method you will use and how you will collect your data (e.g., mailed survey, interviews). Determine the resources that you will require per participant, and then determine the total number of participants that you can afford based on resource constraints. Try for that maximum number, but be aware that the number you choose has no statistical grounding. That is, if you come up short of the number you are not violating any statistical properties of your sample.

V. Summary and Conclusion

This guide has provided information on developing population lists and choosing samples from unknown populations. The fact that a population is unknown or hard to enumerate should not deter you from collecting information. Unknown populations require you to take the first step of generating a reliable population list, rather than relying on a readily available one. We expect that this guidance will assist you in developing such lists. In cases of unknown populations, developing a population list can be a valuable project outcome in itself even if that list is known to be only partial. What is most important in dealing with unknown populations is to analyze and evaluate your data sources at each step. Constant analysis and evaluation will help you understand the limitations of your final product.

Your final list will probably not contain all of the population members and will probably contain entities that are not in your population. We expect that this guidance will help you minimize as much of this over- and under-coverage as possible. Nevertheless, even well-funded statistical sampling projects often rely on less-than-perfect population lists. As we have stressed in this guide, however, the key is to understand the sources of the imperfections. Understanding the imperfections of your list will go a long way to help you understand the imperfections and limitations of data you collect using your list.

⁶⁶ There is no real restriction on how to choose the facilities at this point. You can, in fact, use one of the other informal methods such as a judgement sample. You can also randomly select the facilities from among the “small” and “large” ones.