

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

INEL-95/0076

Test Plan Guidance for Transuranic-Contaminated Arid Landfill Remedial Technology Development

Joe Evans^a
Peter Shaw

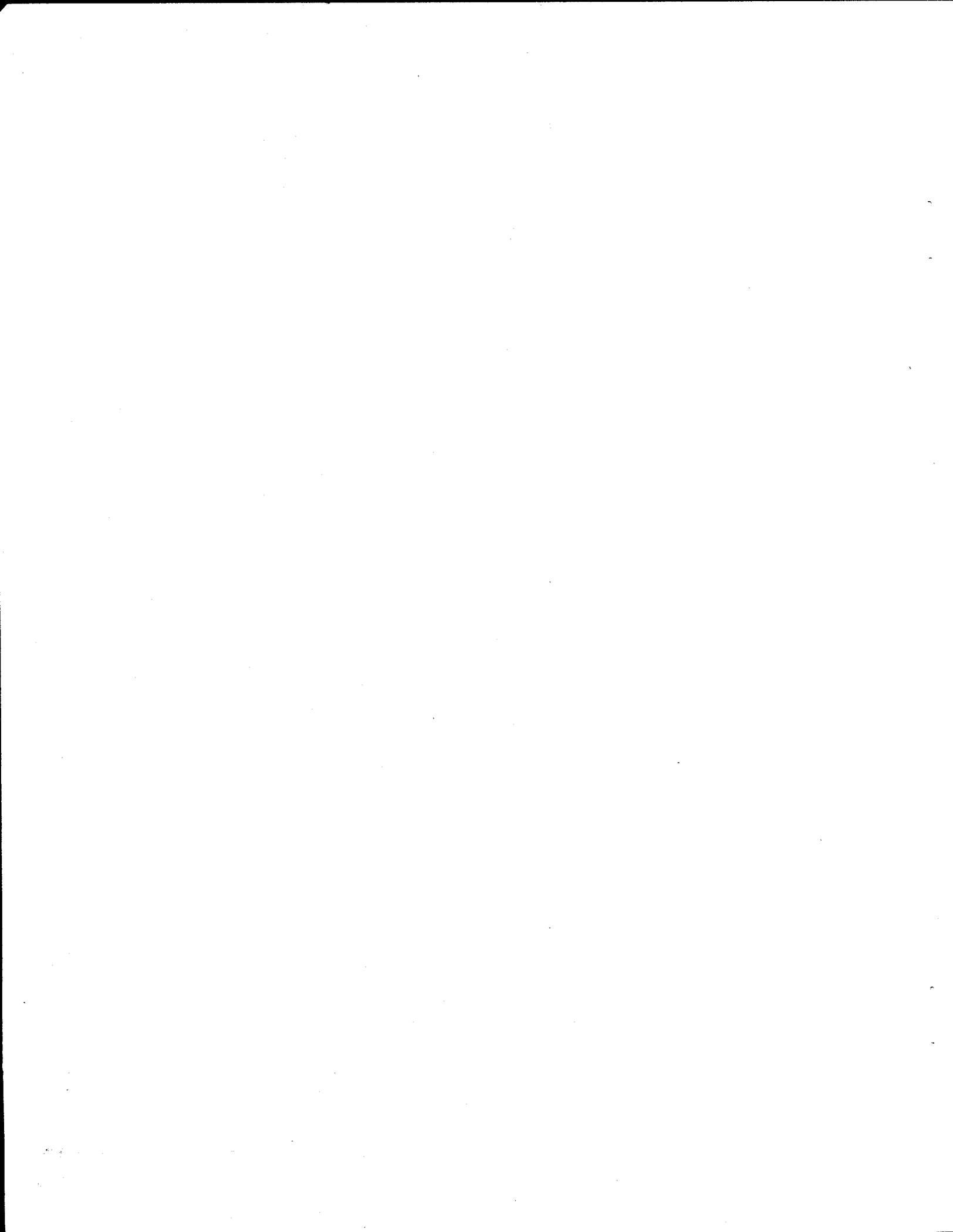
May 1995

Idaho National Engineering Laboratory
Lockheed Idaho Technologies Company
Idaho Falls, Idaho 83415

Prepared for the
U.S. Department of Energy
Assistant Secretary for Environmental Management
Under DOE Idaho Operations Office
Contract DE-AC07-94ID13223

a. Science Applications International Corporation (SAIC).

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
DT MASTER



DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

**Test Plan Guidance
for Transuranic-Contaminated Arid Landfill
Remedial Technology Development**

INEL-95/0076

Prepared by



Joe Evans, SAIC

5/16/95

Date

Reviewed by

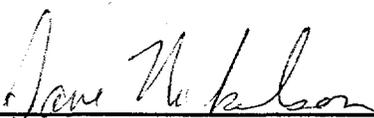


Peter Shaw, Staff Scientist

5/18/95

Date

Approved by



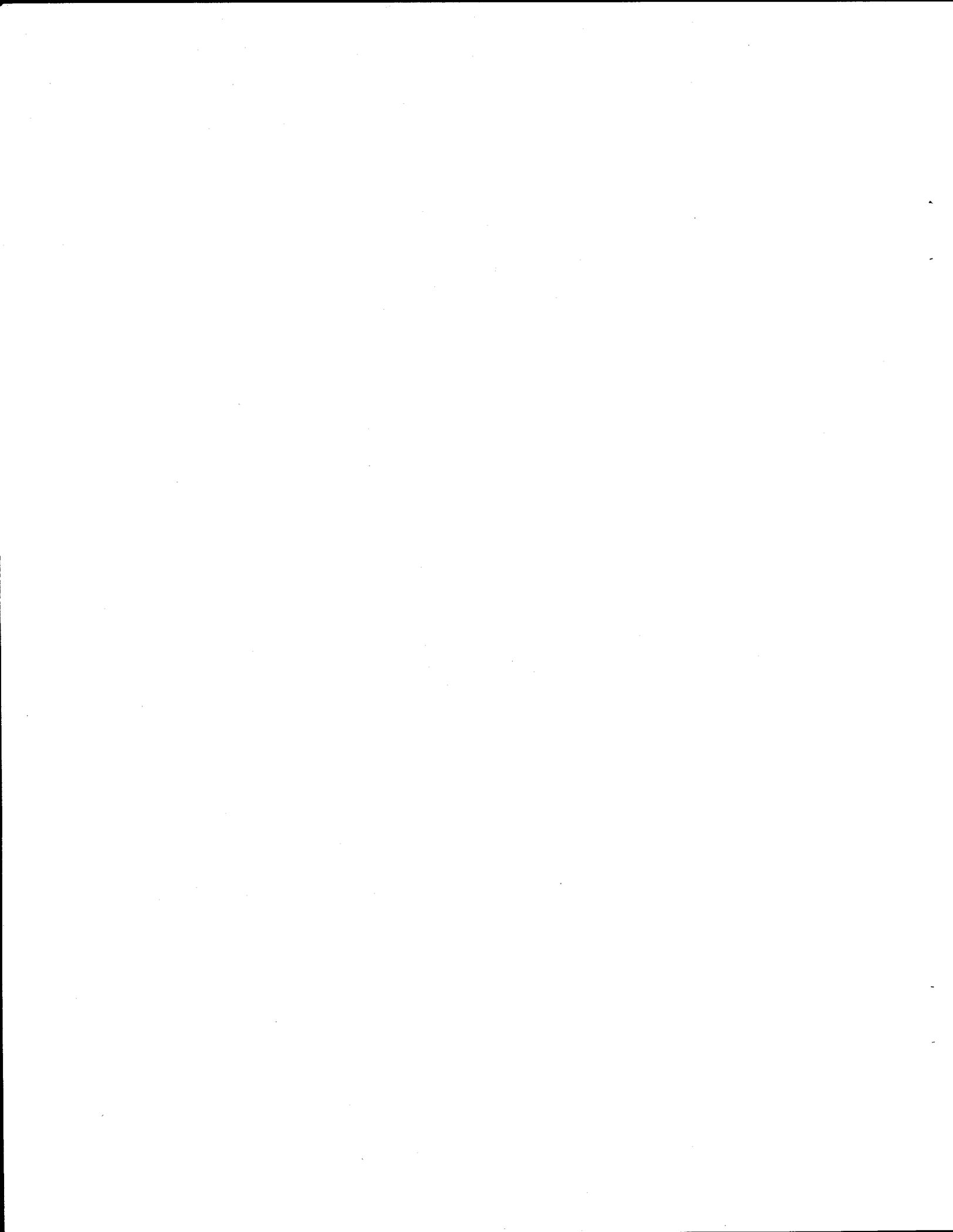
Dave Nickelson, Staff Engineer

5/16/95

Date

ABSTRACT

This document provides guidance for preparing plans to test or demonstrate buried waste assessment or remediation technologies supported by the U.S. Department of Energy's Landfill Stabilization Focus Area, Transuranic-Contaminated Arid Landfill Product Line. This document also provides guidance for development of data quality objectives, along with the necessary data to meet the project objectives. The purpose is to ensure that useful data of known quality are collected to support conclusions associated with the designated demonstration or test. A properly prepared test plan will integrate specific and appropriate objectives with needed measurements to ensure data will reflect the Department of Energy Office of Technology Development's mission, be consistent with Landfill Stabilization Focus Area test goals, and be useful for the Department of Energy Environmental Restoration and Waste Management programs and other potential partners (e.g., commercial concerns). The test plan becomes the planning and working document for the demonstration or test to be conducted ensuring procedures are followed that will allow data of sufficient quality to be collected for comparison and evaluation.



SUMMARY

The U.S. Department of Energy's (DOE) Landfill Stabilization Focus Area (LSFA) supports applied research, development, demonstration, and evaluation of a multitude of advanced technologies dealing with difficult assessment and remediation problems related to radioactive and hazardous buried waste. These innovative technologies are being developed as parts of integrated comprehensive remediation systems for the effective and efficient remediation of buried waste sites throughout the DOE complex. The DOE Office of Technology Development (EM-50) has initiated a number of comprehensive research, development, demonstration, test, and evaluation programs to promote innovative and emerging technologies and technology systems.

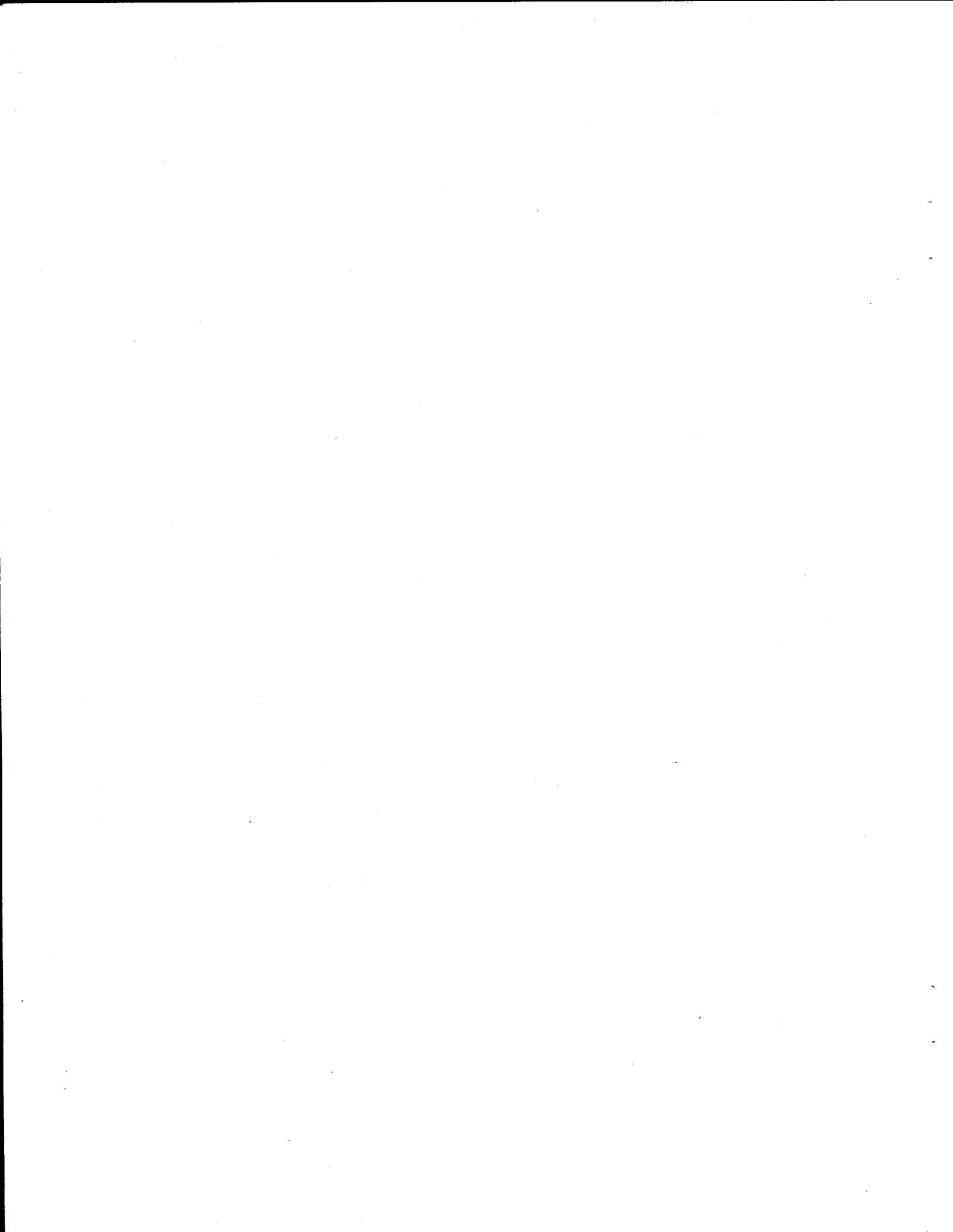
Planning for applied research and demonstration activities, which formerly were sponsored by the Buried Waste Integrated Demonstration (BWID) program, has followed the *Buried Waste Integrated Demonstration Program Technology Test Plan Guidance, Revision 1*. This report is a revision of this BWID test plan guidance. The BWID program activities have been transitioned into the LSFA. This revision discusses data quality objectives (DQOs) and the methodology used to plan development, test, and demonstration activities using a tiered approach. This guidance will be implemented for the DOE LSFA buried waste remediation technologies within the Transuranic-Contaminated Arid Landfill Product Line.

This revision expands the discussion of DQOs and quality assurance (QA). This includes examples for determining test objectives and associated DQOs. The scope of projects this test plan covers includes technologies for assessment and remediation of hazardous and radioactive buried waste. These technologies include site assessment, waste assessment, containment, stabilization, removal, treatment, and disposal. The revisions involve a graded approach to test planning for QA and DQOs and consolidating the quality section similar to the Quality Assurance Project Plan (QAPP) by separating logistics and experimental functions.

The main section of the report outlines the test plan. Appendix A contains a more detailed example of a test plan based on Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) criteria and examples for all types of LSFA technology demonstrations/tests. The following key principles form the basis for guidance:

- Test plans must address the needs of the data user.
- Test plans should follow a graded approach considering the quality of data needed.
- The seven-step DQO process should be followed for all tests.
- The information gathered during testing/demonstration will be used for the remedial investigation/feasibility study (RI/FS) process; therefore, CERCLA criteria shall be addressed.

The test plan guidance is to ensure that the technology development planning and resulting data focus on the waste remediation problems and needs of Environmental Restoration/Waste Management, particularly in obtaining data valuable for effectiveness, implementability, and cost determinations. This test plan guidance is intended to apply to the overall continuum of technology development activities: basic research, applied research, exploratory development, advanced development, engineering development, and prototype demonstration.



CONTENTS

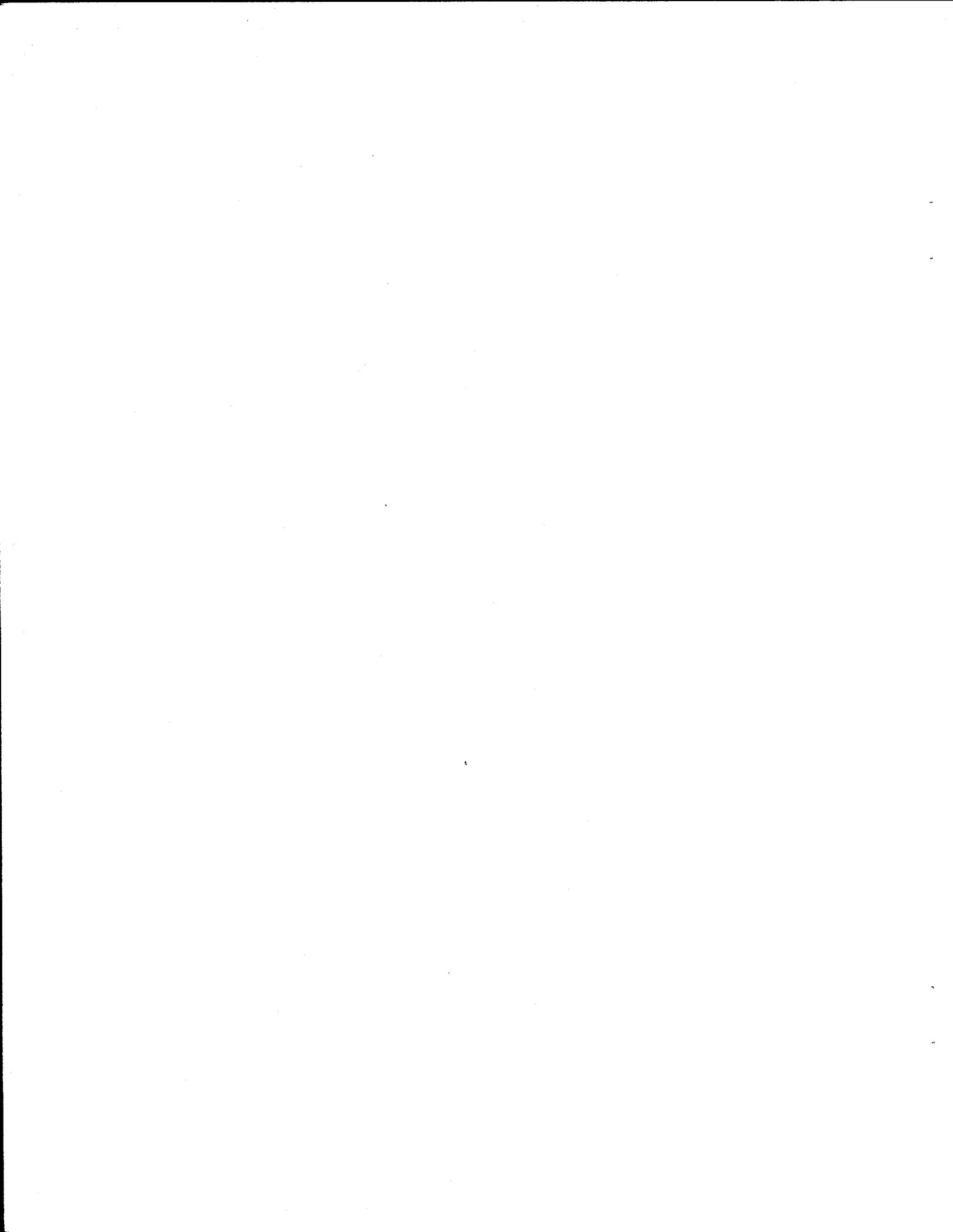
ABSTRACT	iii
SUMMARY	v
INTRODUCTION	1
POLICY	3
INSTRUCTIONS	4
DEFINITIONS	6
REFERENCES	11
Appendix A, Sample Test Plan Guidance	A-1
Appendix B, Test Logistics and Deployment Guidance	B-1

FIGURES

A-1. Development process for data quality objectives.	A-6
--	-----

TABLES

A-1. Full-scale operation estimates.	A-8
A-2. Nine CERCLA criteria.	A-16



Test Plan Guidance for Transuranic-Contaminated Arid Landfill Remedial Technology Development

INTRODUCTION

This technology test plan and data quality objective (DQO) guidance provides instructions for preparing test plans and DQOs to test or demonstrate buried waste technologies supported by the Landfill Stabilization Focus Area (LSFA), Transuranic-Contaminated Arid Landfill Product Line. This document provides guidance for development of DQOs along with the necessary data to meet project objectives. Development of the project DQOs using the process identified in this guidance should be an integral part of the test plan.

The purpose of the technology test plan is to ensure that useful data and data of known quality are collected to support conclusions associated with the designated demonstration or test. A properly prepared test plan will integrate specific and appropriate objectives with needed measurements to ensure data will meet requirements of the data user as defined below.

The technology test plan should be written to reflect the Department of Energy (DOE) Office of Technology Development's mission, to be consistent with LSFA test goals, and to ensure the data collected by the test are useful for the DOE environmental restoration and waste management programs and other potential partners (e.g., commercial concerns). The test plan becomes the planning and working document for the demonstration or test to be conducted. It is not only used for preparation and planning of the demonstration, but it is used during the demonstration to ensure procedures are followed that will allow data of sufficient quality to be collected for comparison and evaluation.

The following key principles form the basis for guidance:

- Test plans must address the needs of the data user. Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) criteria are key in determining these needs. Clients (e.g., Environmental Restoration) are included in preparation and planning. This is an integral part of the DQO process. Test objectives must be directed to these agreed upon needs.
- Test plans should focus on a graded approach. Consider the quality of data needed based on customer necessity. The DQO process presented in this document should be followed and applied appropriately. It is not necessary to include rigorous DQOs for all test objectives.

An example of what constitutes critical and noncritical objectives has been included in the definition of test objectives. This example helps to specify when DQOs need to be more rigorous based upon the determination to be made. Projects requiring lower quality data, however, still need defined objectives and specific measurements, even if the quality assurance (QA) associated with these measurements is less rigorous.

- The seven-step DQO process (see Definitions for this process) should be followed

for all tests. This is outlined in detail in the guidance presented, along with examples of how to incorporate DQOs into the test plan.

- Emphasis needs to be placed on critical project objectives with associated measurements. Principal investigators in concert with program management must separate critical objectives from noncritical objectives based on environmental restoration needs and the scope and nature of the project. This includes determining quantitative versus qualitative objectives and associated measurements.
- The information gathered during testing/demonstration will be used for the remedial investigation/feasibility study (RI/FS) process; therefore, CERCLA criteria shall be addressed.

POLICY

All test and demonstration data collection activity associated with technology development for the LSFA Transuranic-Contaminated Arid Landfill Product Line shall be covered by an approved test plan. The draft test plan shall be submitted to and approved by the LSFA Transuranic-Contaminated Arid Landfill Product Line manager. Personnel with technical expertise in the technology being demonstrated, expertise in safety, quality assurance, regulatory requirements, and operations for the host site demonstration, will be utilized as required for reviewing the draft test plan. The test plan must be approved prior to commencing any test or field activities.

INSTRUCTIONS

Test plans should be developed for a complete test of a technology or system. The test plan shall include the following sections and appendix:

ABSTRACT

ACRONYMS

1. PROJECT DESCRIPTION
2. ORGANIZATION AND RESPONSIBILITIES
3. TEST OBJECTIVES
4. QUALITY ASSURANCE OBJECTIVES FOR CRITICAL MEASUREMENTS
5. SAMPLING PROCEDURES AND DATA COLLECTION
6. DOCUMENT CONTROL
7. ANALYTICAL METHODS, CALIBRATION PROCEDURES, AND LABORATORY QUALITY CONTROL
8. DATA REDUCTION, VALIDATION, AND VERIFICATION
9. QUALITY ASSURANCE
10. REFERENCES

APPENDIX A

- 1 Test Logistics and Deployment
- 2 Health and Safety
- 3 Residuals Management
- 4 Procedures

All sections must be addressed, including the mandatory appendix, even though sections may be simplified for smaller activities or laboratory activities. If any of the sections are not applicable, the heading shall be included with the statement, "This section is not applicable." Include a statement that justifies why the section is not applicable. For example, in a proof-of-principle test there might be no laboratory data generation (Section 7) planned. A statement to this effect would be made with a justification for this based upon the nature of the test. Specific instruction from the customer is excellent in this regard, especially if no quantitative data are to be generated. A statement might also be included as to how qualitative data will replace laboratory data.

Appendix A goes through each section and describes the test plan further. The test plan appendix is mandatory but is not part of the formal DQO process. Following health and safety guidelines is assumed and is mandatory but will not guarantee quality results or satisfactory data. However, in proof-of-principle tests, for example, how safety issues are addressed and observations of the system's safety might be part of the qualitative data generated.

DEFINITIONS

Test Objectives: Test objectives are determined in order to evaluate a process or technology. Test objectives can be quantitative and qualitative, usually corresponding to critical and noncritical. However in some tests the qualitative objectives may be the critical objectives (see **Critical and Noncritical** definition below). Quantitative statements are supported by statistically based comparable numerical data. Qualitative statements are more subjective and are often based on observational measurements. Quantitative or qualitative objectives may be used for evaluation to determine the success of a technology, though quantitative is preferred, particularly in comparing technology demonstrations at different sites. Quantitative test objectives can be formalized with DQOs, which require measurements with defined QA requirements. These measurements then result in producing comparable data useful for those interested in technology selection and application.

When demonstrating a technology or determining whether a process is suitable for further study, a researcher should develop objectives that will demonstrate the success or failure of the process being tested. Regardless of whether the actual process or technology is a failure, does not mean the test was a failure. The determination of whether a test is successful is based upon whether conclusions can be drawn from the data. This means that if the data show, with some degree of confidence, that the chosen process or technology is unsuitable for further study, then the test is successful. Chosen objectives should be considered carefully so that positive or negative results will lead to definitive conclusions. Collected data must also have a degree of confidence associated with the determination made so that decisions can be based upon reasonable evaluations.

One data point, for example, does not give any degree of confidence with the associated measurement. Three data points may give some reasonable range with which to evaluate a process, but if any one of these points are questionable, due to collection error or variability, then there is often little or no confidence in the collected data to properly evaluate a technology or process. Some good rules of thumb are to remember that, statistically, seven data points will often reduce the confidence range such that definitive decisions can be made if collected data are compared against another process, and 30 data points is usually enough to define a population.

When designing an evaluation test, decide how the data will be used to determine if the information can be considered critical or noncritical. Critical data should directly evaluate the success or failure of a process or technology. An example of this type of determination is presented below.

In evaluating the success of the cryobarrier technology (underground frozen wall created to reduce mobility of contamination), it was important to determine if a frozen barrier or wall could be created beneath the surface and specifically if, in dry soil conditions, soil moisture could be increased in this wall to create an actual barrier to contamination mobility. In this example, two types of measurements could be made for determining technology success. These measurements are chosen to illustrate how critical and noncritical objectives can be chosen, but the measurements are not meant to be the only examples of measurements to evaluate this technology.

Within the limits of this frozen barrier, several temperature probes can be placed to determine if in fact the soil temperature is reduced to a specific objective. It is not feasible, however, to place enough probes underground to measure the temperature of every square foot

within the confines of this frozen wall of water. The wall can therefore be divided into a defined number of cubic foot sections, and several temperature measurements can be made at randomly chosen locations. In this manner, a statistical confidence can be determined on whether this process was successful when the data from these measurements are collected.

If, for example, 30 data points are randomly chosen and all 30 indicate the desired temperature was reached, then the process can be determined successful in achieving the stated objective with a statistical degree of confidence. If, however, a few of these data points (e.g., two or more) show temperatures of the wall did not meet stated objectives, then it is likely that not only did these measured areas show holes or leaks within the barrier but there are also other areas within the barrier not meeting stated objectives. These measurements and the associated objectives could be considered critical for technology evaluation. Positive or negative results would indicate whether the technology was a success or a failure.

In addition, it may be beneficial to place a moisture measuring device on the outside of the wall where leaks may be expected to occur. Assuming that it would be impractical to make these measurements at all points surrounding the wall, a few select points may be chosen. Indications of increased moisture outside the barrier would also indicate that the process was not working as expected.

Because, however, there are many more possible data points that could not be measured outside the defined barrier, this measurement would not be as definitive as the temperature measurement of the wall. Therefore, if every probe placed outside the wall (e.g., a defined number like 30) showed no increase in moisture, then this does not mean the wall is working as expected because there could be undetected leaks. In this case, a negative result (all 30 data points showing no increase in moisture) would be inconclusive. This measurement should therefore not be considered critical. Conclusions could be drawn with positive results but not with negative results. It may be a worthwhile measurement, however, and could be considered a noncritical measurement that would supply supporting data.

The example presented above contained quantitative objectives that could lead to definitive conclusions. In the example of the moisture measurements, however, as previously noted, not always could conclusive results be obtained. This is similar to many types of qualitative determinations. They probably can't give definitive determinations if they stand alone but may be able to be combined with other measurements of a similar nature to increase the probability that the process or technology is or is not working as expected.

For example, along with the moisture measurements made above, some determination of the amount of water added to the soil could assist in determining whether there were leaks in the wall. It would be expected that a definitive amount of water could be added to create this frozen barrier of given dimensions. If this expected result was taken together with the moisture measurements, and the moisture measurements showed no leaks, then this would be further evidence that the technology was working as expected. If, however, the amount of water added passed what would be theoretically expected, then it may be determined that there were leaks in the barrier.

This measurement, however, would be more qualitative in nature because there would be several uncertainties associated with the amount of water this frozen barrier may be expected to hold. This determination could therefore not be used alone to determine the success of the technology but could be combined with other data to give additional evidence and increased

confidence about whether the technology was a success.

Test objectives, however, are sometimes qualitative and may be based on subjective observations. As an example, when demonstrating in situ vitrification for the Superfund Innovative Technology Evaluation (SITE) program, many critical objectives were nonquantitative because of the nature of the chosen site. Destruction and removal efficiencies, which were normally a critical part of the evaluation for this technology, could not be accomplished because contamination at the site was too low. It was important, however, to evaluate whether this technology could be successfully demonstrated on a pilot scale. Could the technology be operated safely without significant hazards? Was it successful in creating appropriate melts? These were test objectives that could only be determined subjectively. Quantitative DQOs for this technology included such items as air emission measurements. These could be determined quantitatively, and reported results could be directly compared to other technologies.

If test objectives are all qualitative, then no standard comparable measure of the technology will be available. Principal investigators and project managers must consider the consequences associated with noncomparable data if only subjective observations are to be made. The value of performing these types of demonstrations should be carefully considered.

Successful test plan formulation and ultimately test execution depends on a close association between the technology development group and the environmental restoration customers. After obtaining a clear understanding of customer requirements, the principal investigator and project manager have the final responsibility for determining the measurements to be taken and the quality needed for these measurements based on test objectives and funding.

Critical and Noncritical Objectives: Critical objectives are key to the decisions to be made (i.e., evaluating the technology) based on the results obtained. Data obtained for these objectives will be used to determine the direct success of the test or demonstration. Noncritical objectives are additional pieces of information concerning technology performance but may not be critical for the overall evaluation or may not be comparable to other similar applications. All tests should have separate critical and noncritical objectives. There may be fewer critical objectives for proof-of-principle type applications and these objectives may or may not be quantitative measurements. The DQO process should be followed for all objectives. All reviewer's (both environmental restoration and technology development) should especially be aware of three types of test objective problems, particularly with critical objectives.

1. Objectives that are poorly defined—data quality will not satisfy the stated objectives
2. Objectives that are missing in the test design—data will not be taken for critical issues
3. Objectives that are unnecessary—extra data will be taken incurring excessive cost or delaying test.

After the reviewer identifies these situations, the principal investigator and project manager must then decide if the test or test objectives need modification, factoring budget and schedule considerations with technical review comments.

Data Quality Objective: Data quality objectives (DQOs) match test objectives and project requirements with QA objectives (see below) to ensure that collected data will meet previously

specified test objectives. DQOs should be thought of more in terms of a process and not finalized objectives. Not all test objectives are considered DQOs, but all DQOs should be test objectives. Details regarding the development of DQOs, along with examples, are presented in Appendix A.

DQO Process:

1. State the Problem
2. Identify the Decision
3. Identify Inputs to the Decision
4. Define the Study Boundaries
5. Define the Quality Based on Customer Needs for the Project
6. Specify Limits on Decision Error
7. Optimize Design for Obtaining Data.

The DQO process should be thought of as a formalized test design process to ensure data quality requirements will be realized.

Quality Assurance Objectives: These are quality indicators for measurements used to determine quantitative DQOs. They include precision, accuracy, completeness, comparability, and representativeness. For example, a quality assurance objective might be to take a measurement repetitively three times to obtain precision or compare a measurement with a known standard for accuracy. In order to ensure collected data are useful, it is often necessary to collect a statistically significant quantity of data points. A statistician might prove to be useful in this regard.

Project Categories: In determining how important a project should be and subsequently the amount of preparation required and quality of data needed, a graded or category approach has been used by the Environmental Protection Agency (EPA) in the development of test plans. It should be noted that because a project is considered to be of a lower category, this does not mean the DQO process can be ignored or even lessened. The DQO process is a means of helping to determine the amount of quality assurance required and how the data will be used, thereby defining the overall critical nature of the investigation. This is probably a point of confusion for many investigators. Most often, investigators will consider the DQO process only to be necessary for high-level projects. This assumption, however, is false! In fact, it is the process itself that helps define the level of quality assurance required.

Currently there is no specific guidance within DOE on how to define different levels of quality for different projects. The quality of data required for a specific project is left to the discretion of the principal investigator. If, however, the principal investigator does not take into consideration the needs of the customer (as outlined later in this document), then the project will most likely fail, as the data quality will often not satisfy customer needs and the data will not serve its purpose. Therefore, while some guidance may be given concerning specific levels of QA, it should not be considered definitive until specifically defined within DOE.

Even when using a category approach, it should be remembered that these are EPA definitions and may not be applicable to all projects. In addition, EPA has not set this category approach as a standard but instead has allowed it to exist as overall guidance.

Four separate categories of research projects have been defined. The most stringent and highest level is for projects directly related to "regulatory compliance." These projects require that all aspects of quality assurance be considered and that only standard methods of measurement be used. The data from these types of projects must be able to stand up in a court of law. Very few research endeavors fall into this category.

The next level is considered as "technology acceptance." This is the category used for the SITE program. The purpose of this type of test is to be able to state that a particular process is able to work or not work for a specific site based upon one demonstration. Data must therefore be thoroughly quality controlled, usually only standard, accepted methods of measurement are used (method verification is required for nonstandard methods), and supplemental data are taken to ensure a high level of confidence in the results.

The next level down could be considered "treatability testing." These results are often combined with results of other tests to determine the feasibility of a particular process or technology. This type of test receives less funding and, subsequently, testing is not as rigorous.

The lowest level or category considered by EPA is "proof-of-principle" type of testing. Often this may only be used to determine if research in this area should go forward, and therefore minimal quality control is needed.

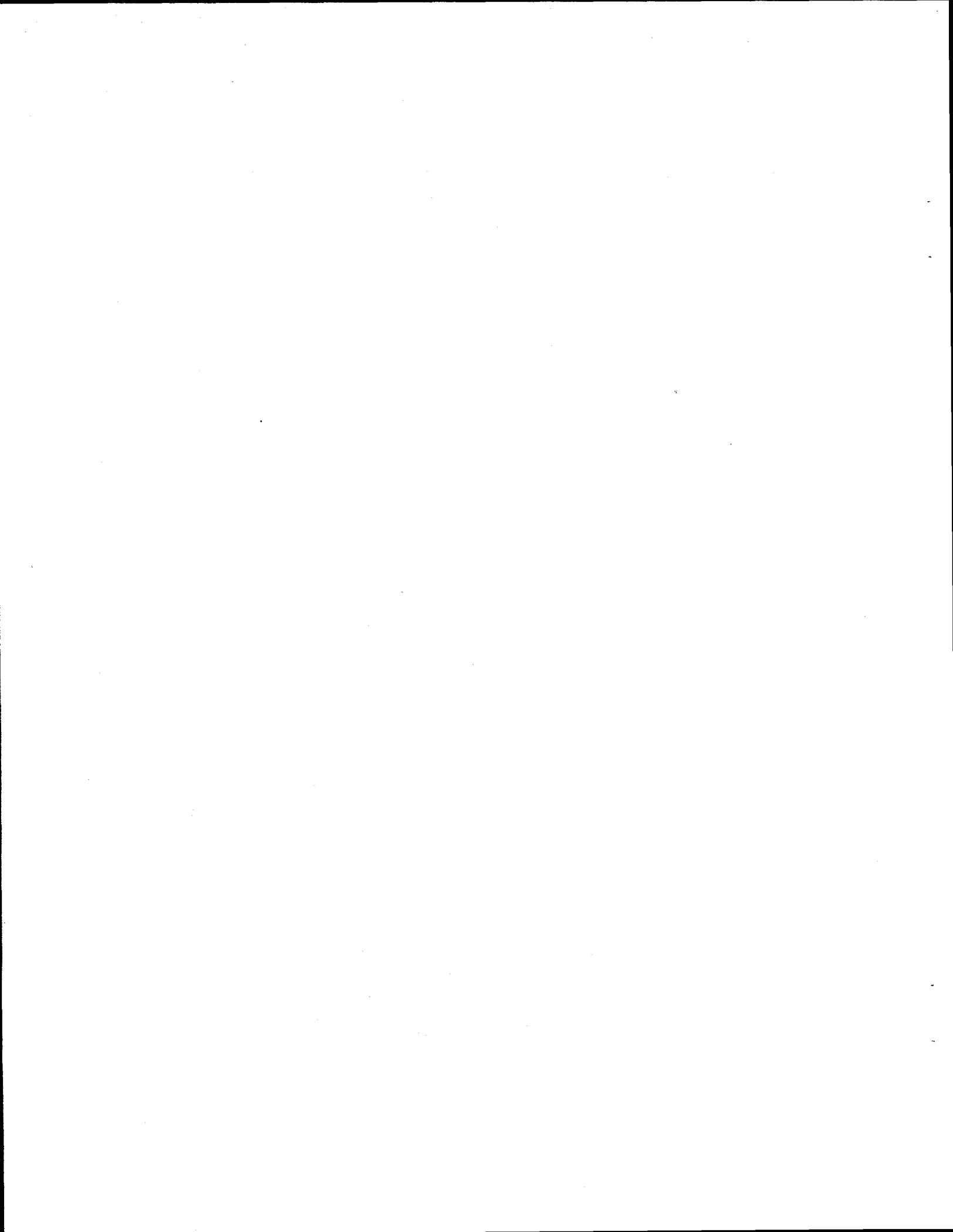
When determining what level of testing a project may fall into, the principal investigator should remember that quality assurance is the same for all tests as defined by the DQO process. Most importantly, this process determines the data quality needed. Quality control will vary based upon the category under which the project is best defined.

As additional guidance for the principal investigator, some EPA definitions have been presented to aid in determining project category. Guidance concerning the quality control required for each of these different types of projects can be obtained from the EPA. (Note reference number 4) The purpose of this document does not encompass the task of defining quality control for each specific QA category. The EPA has prepared separate documents to cover guidance for Category I, II, III, and IV projects.

REFERENCES

1. INEL Environmental Restoration Program, *Program Directives*, 5.10, "Test Plans," April 2, 1993.
2. EG&G Idaho, *Quality Manual*, QP-11, "Test Control."
3. DOE Order 5700.6C, "Quality Assurance."
4. U.S. Environmental Protection Agency, *Preparation Aids for the Development of Category II Quality Assurance Project Plans*, February 1991, EPA/600/8-91/004.
5. U.S. Environmental Protection Agency, *Guidance for the Data Quality Objective Process*, EPA QA/G-4, Final, September 1994.
6. U.S. Environmental Protection Agency, *EPA Requirements for Quality Assurance Project Plans for Environmental Data Operations*, EPA-QA/R-5, Interim Final, August 1994.
7. EG&G Idaho, Inc., *Buried Waste Integrated Demonstration Technology Test Plan Guidance*, Rev. 1, EGG-WTD-9800, January 1994.

Appendix A
Sample Test Plan Guidance



Appendix A

Sample Test Plan Guidance

1. PROJECT DESCRIPTION

1.1 Technology Agreement

Define the technology agreement with the funding source. This may be as simple as identifying the technical task plan (TTP), memorandum of understanding (MOU), or other cooperative agreement by which the demonstration was developed.

1.2 Technology Description and Background

Include information necessary to permit a technical reviewer, who is unfamiliar with the project, to assess test objectives and data quality including data quality objectives (DQOs) and sample strategy.

Examples of information to be included are a site description, a technical background of the technology, results of predemonstration data or preliminary investigations, if available, and the purpose for which the technology is being demonstrated. Include diagrams, flow charts, and process descriptions, as necessary. Include illustrations to assist in understanding the test.

1.3 Test Schedule

Describe the order in which activities will occur, detailed schedule of the test, and its associated activities/analysis. Indicate milestones and decision points.

1.4 Contingency Plan(s)

State the contingency plan(s). Explain how implementing the contingency plan(s) would affect test and quality objectives, schedule, cost, and possibly location. Contingency plans may be presented in a logic diagram format.

2. ORGANIZATION AND RESPONSIBILITIES

The following organization and responsibility information is required:

- Name all key individuals in charge of every major activity in your project and describe their activities. Examples include the project manager, principal investigator, quality assurance (QA) officer or coordinator, field manager, and subcontractor project manager, if used. The customer such as the name of the ER representative should be given here. Their role in the review test plan or attendance at demonstration activities might also be included.
- Show a chart defining organizational responsibilities; include subcontractors, if used.
- Describe the responsibilities of the organization conducting the test and the delegation of authority for conducting the test.
- Indicate personnel support requirements, including monitoring and surveillance, as needed.
- List the required test personnel qualifications and level of training.
- Show the QA organization. QA should be independent.
- Show the health and safety organization. This organization should also be independent.

3. TEST OBJECTIVES

State the purpose of the test, type of information the test will furnish, and the DQOs. As a minimum, test objectives will be drawn from the current version of the *BWID Test Goals and Objectives* (EGG-WTD-11112).

The DQO process is a formalized mechanism to optimize test design. Following the DQO process ensures that the data quality needed to evaluate the critical objectives can be achieved. Some level of data quality is always needed regardless of the test being performed. Unfortunately, data quality may often be considered outside the scope of the project. It is, however, always a key component of collected data. It may be "loosely" defined, but must be considered when determining test objectives. Figure A-1 is an outline of the development process for DQOs.

Included below is the seven-step DQO process. This process should be followed for all demonstrations or tests of a technology. Not all tests, however, will require formalized objectives with quantitative measurements.

1. State the Problem (define the purpose of the test)

Determine what will be evaluated. Focus on the primary issue. If the primary purpose is to determine whether a technology will be successful, state specifically what constitutes success. Is this a specified reduction efficiency or is the problem to evaluate a capability such as, "can the process be deployed successfully at a full-scale level?" This type of problem needs to be defined in terms of specific objectives.

The primary objective of the test may be quantitative or qualitative. (Refer to previous definition on "test objectives.") This decision must be made by the principal investigator. If the purpose of the test is qualitative, consider the effort and cost associated with making a qualitative statement to evaluate the technology performance and decide if it's worth the time and expense.

Qualitative objectives and subsequent measurements do not produce easily comparable data. Qualitative objectives, however, may be the primary determination for evaluation and include subjective measurements. An example of a qualitative objective critical for evaluation of a technology could be simply to decide if field deployment can be achieved. This may be determined by a number of subjective measurements such as operability of equipment, required personnel, or a determination of feasibility of the process.

Quantitative statements provide direct measures of comparison. These are usually developed into formalized DQOs. Examples of quantitative objectives are included at the end of this section. In some situations, while an objective may seem qualitative it is actually quantitative. It could be as simple as the number of bucket loads required to empty a pit. This still transfers into a quantitative measurement with an associated precision and accuracy. Determination of a full bucket and the number of buckets required provide a measure of comparison.

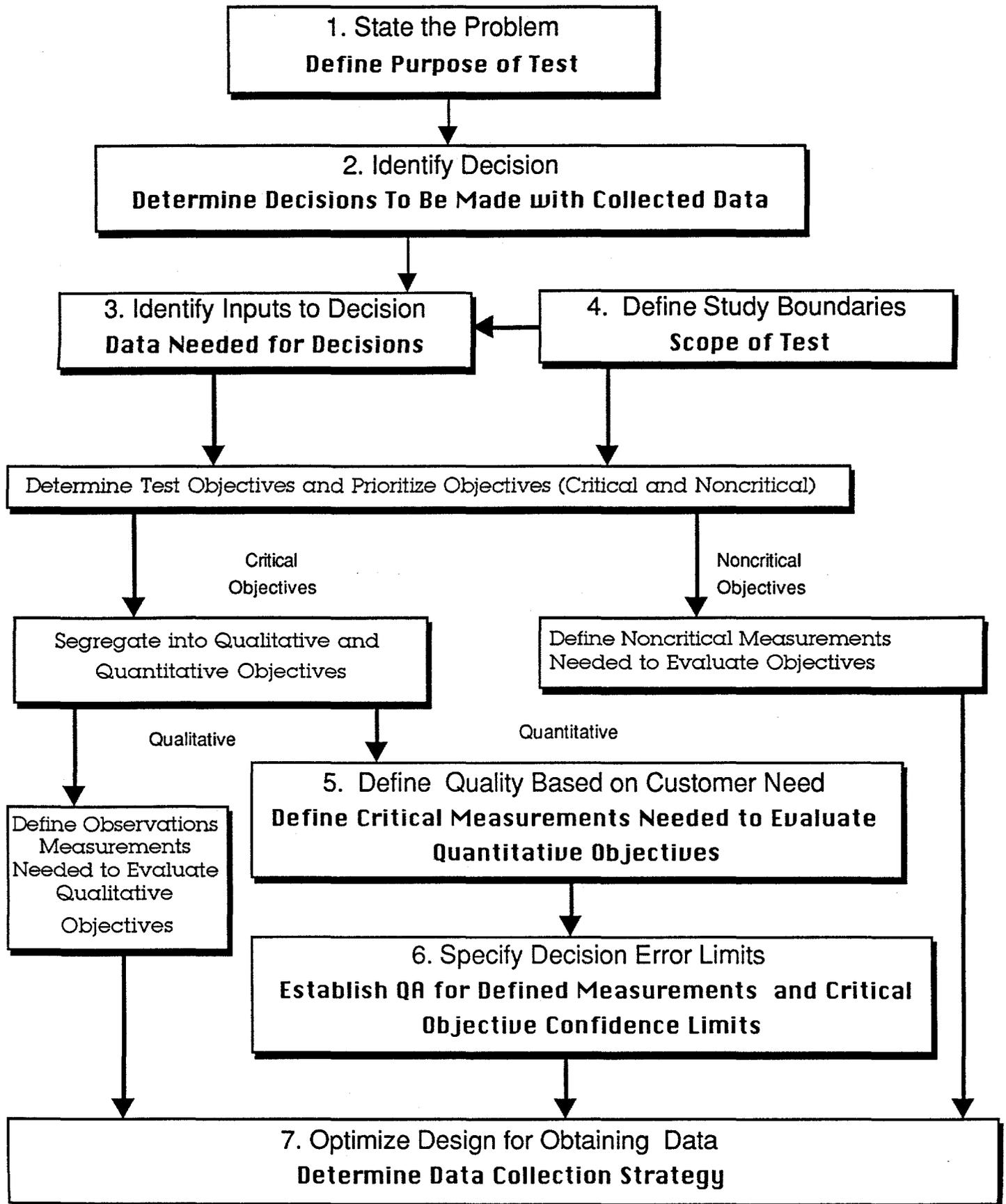


Figure A-1. Development Process for Data Quality Objectives

2. Identify the Decision (determine decisions to be made with collected data)

Determine the decision to be made based on performance results and how the results will be used. Consider the needs of the client (waste management or environmental restoration). Will the decision be important for regulatory purposes? Will the decision be to determine if this technology can be used for a particular application? Will the decision to use this technology be based on performance compared to other technologies? Is it important to determine if the tested technology is ready for commercial application or simply to determine if additional tests of the technology are warranted?

3. Identify Inputs to the Decision (data needed for decisions)

What data will be collected for evaluation of the technology and subsequently input into the decisionmaking process?

NOTE: Sufficient data should be collected by the test to address as many of the parameters in Table A-1 as applicable, commensurate with the tier of the project. For example, if the test is a laboratory study of one aspect of an entire system, many of the remediation full-scale items such as time to remediate (e), remediation volume (f), or total number of assemblies (j) might not be applicable. The information requested in Table A-1 will be reported in the technology evaluation report. This information will be used to evaluate the technology for full-scale operation or to assess whether further development is viable.

Quantitative test objectives should be made into DQOs. Indicators of DQOs are precision, accuracy, representativeness, completeness, and comparability, as determined by QA objectives defined in Section 4. If a quantitative evaluation will be made based on test results, then outlined in the following paragraphs is a step-by-step process for better defining project objectives and producing DQOs. Subsequent steps of the DQO process are highlighted:

- List the performance objectives important to the evaluation of the technology.
- Segregate objectives into qualitative and quantitative statements.
- Segregate quantitative objectives into critical and noncritical objectives and prioritize. Critical objectives are key to the decisions to be made based on the results obtained. Noncritical objectives are additional pieces of information concerning technology performance but may not be critical for the overall evaluation or may not be comparable to other similar applications.

The number of critical objectives is based on project scope, cost, and customer requirements. Quantitative critical objectives will require the most accurate and precise measurements. Too many critical objectives might divert QA resources better spent on a limited more manageable number. Often a project will have no more than about five critical objectives, even fewer if it is at a demonstration level. A proof-of-principle

Table A-1. Full-scale operation estimates.

- a. Operation and maintenance cost for the life cycle of the technology demonstrated. Operation and maintenance cost includes storage and disposition of waste, operation of equipment (manpower, electrical power, etc.) and maintenance (cost of replacement parts, manpower, etc.), and decommissioning.
 - b. Acquisition cost, including capital.
 - c. Volume (m³) of input material, toxicity level, mobility, and physical state.
 - d. Volume (m³) of contaminated material produced (gloves, bags, tools, process water, etc.) and physical state.
 - e. Time to remediate the input volume.
 - f. Time to remediate 50% of the input volume.
 - g. The process output, their waste classification, concentration or activity, toxicity level, mobility, and volume (m³).
 - h. Time until (a) line item and capital equipment replacement costs exceed 50% of acquisition cost and (b) line item and capital equipment repair and maintenance costs are estimated to approach 50% of acquisition costs. Compare to expected useful life (see item a above).
 - i. Assemblies and subassemblies specified in the design and functional operating requirements, and longest procurement lead times for custom and engineered components.
 - j. Total number of custom assemblies and subassemblies.
 - k. Time for scheduled maintenance.
 - l. Time for unscheduled maintenance.
 - m. Time from now to field demonstration. Now is considered to be the approval date of the technical evaluation report.
 - n. Time from now to full-scale operation.
 - o. Mean years of specialized training required for full-scale operation (academic and facility specific).
 - p. Labor hours per hour of operation.
 - q. Time for the replacement parts to exceed 50% of the acquisition costs.
 - r. The list of permits required for a full-scale operation.
 - s. Time to construct or fabricate the equipment necessary to operate the process.
 - t. Types and quantities of utilities (water, electricity, etc.) required during the process.
 - u. Time to train technical personnel and operators.
 - v. Extra time and costs prudently estimated to cover contingencies (i.e., unplanned events outside the control of the principal investigator).
-

type of project may have only one or two critical objectives based upon the fact that a smaller budget is available.

4. Define the Study Boundaries (scope of test)

- Eliminate objectives that the project budget cannot afford to measure. If the primary objective is to determine if process emissions meet a specified regulatory requirement, then this objective will require the focus of the project's budget. Determine the amount of time and effort required to evaluate critical project objectives and eliminate, as necessary, those objectives that will not enhance overall conclusions concerning the technology.

It is critical to have the test plan reviewed to determine if the project can achieve the specified objectives. Many principal investigators will proceed simply because money is available and the project has been determined worthwhile. If, however, the project budget does not allow for enough, or the right type of measurements to be made to achieve the defined objective, then a decision should be made on whether to proceed or if other objectives may present a worthwhile goal. This must be a collective decision made by the QA reviewer, the principal investigator, and other key personnel and should be based upon the needs of the customer.

- Define the scope of the demonstration.

List the number of tests, test location, test duration, and schedule.

Define the subject area of the test and provide a brief description. Define and justify the data collection strategy for the critical measurements as defined by the critical objectives.

For example, if determining the amount of dust generated by the process is a critical objective, then the data collection strategy (why is it critical to collect these data, purpose of collected information, control samples to be collected for comparison, description of the overall sampling and test design) should be included.

5. Define the Quality Based on Customer Needs for the Project (define critical measurements needed to evaluate quantitative objectives)

Ask yourself what the client needs are. With what data will these data be compared and what are the QA objectives required? Determine if the project is a full-scale demonstration, a bench test or applied research, or simply a determination for a proof of principle. The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) criteria are additional pieces of information defining the needs of the customer. These criteria must be included defining the project objectives and determining how the data will be used. It is the principal investigator's responsibility to ensure these needs are satisfied.

State the customer need that the technology is satisfying. Describe how the testing

meets the customer need. This involves interaction with the data users. The data users should play a key role when initiating the DQO process.

A project needs to be defined in terms of how the information gained from the test will be used. Should the data stand alone or will it be combined with data from other similar demonstrations or tests? Are the data to be used to perform additional testing or are the data to be used to determine commercial implementation? Answers to these questions require different data objectives. Therefore, the test may be defined as proof-of-concept, applied research, or full-scale demonstration. Once the QA objectives are determined, appropriate decisions can be made concerning test objectives and subsequent measurements.

6. Specify Limits on Decision Errors (establish QA for defined measurements and critical objective confidence limits)

Write critical objectives in terms of confidence limits and include false positive and false negative error possibilities. Determine the risk you are willing to accept with respect to making a decision error. What are the consequences of incorrect decisions or conclusions based on these results?

For example, if you falsely conclude a technology is successful based on collected data, when in fact it did not achieve the stated objective(s), what will be the result? Is it likely more tests will be performed or will there be immediate acceptance of this technology for implementation purposes? If the technology is scheduled for implementation based on a given demonstration or a single test, then decision error limits will need to be very conservative. This then requires an increased level of quality data. Statistical evaluations may be required based on sampling and analytical variability and the number of samples collected. (See examples of statistical confidence limits associated with quantitative objectives in the example below.)

The following questions should be considered. Is it important to be accurate and precise in citing the critical objectives? Do you want to quantitatively evaluate a technology? What is important for the evaluation?

7. Optimize Design for Obtaining Data (determine data collection strategy)

Repeat the six-step DQO process listed above for each stated objective. This will aid in optimizing the test design and critical objectives.

Presented below are examples of quantitative project objectives:

- Evaluate dust generation of the process being measured.

Poor statement: This objective puts no limits by which you can measure success or failure. It does not allow the opportunity to establish data quality for the measurements being taken.
- Evaluate if dust levels generated by the process meet regulatory requirements of less than $x \text{ mg/m}^3$.

Better statement: This objective defines a specific success or failure associated with the results of the test being conducted. It is now assumed that the defined regulatory level sets a standard that should be met during the test. It also allows for specific quantitative measurements and a minimum level of data quality. QA objectives can be defined.

- Determine at a 95% confidence level if dust levels generated by the process meet regulatory requirements of less than $x \text{ mg/m}^3$.

Best statement: Adding a confidence level to the objective now sets a specific data quality needed with the measurements used to determine the stated objective.

Define critical and noncritical measurements associated with the critical and noncritical objectives. Define the key input and output parameters measured and the test conditions. Critical measurements will be those measurements required to evaluate the primary objectives and will have associated QA objectives as defined in Section 3. Noncritical measurements are those measurements associated with secondary objectives and whose QA objectives are more loosely defined.

For example, when evaluating a grouting technology, it was important to determine the reduction in bedrock porosity within the grout curtain. Packer tests, used to determine in situ porosity, were considered critical measurements and required specific QA objectives in terms of accuracy and precision. Noncritical measurements included slug tests, which were also a measure of in situ porosity, but were considered only as secondary information. The level of QA was much less for these measurements.

4. QUALITY ASSURANCE OBJECTIVES FOR CRITICAL MEASUREMENTS

Quality assurance objectives are specifications that measurements must meet in order to achieve project objectives. QA objectives are needed for all critical measurements. Once QA objectives have been established as part of the DQO process, the measurement systems are then designed to meet them.

4.1 Determining QA Objectives

Once critical measurements have been established based on defined objectives, QA objectives define the quality of the needed measurement. The quality of the measurement is defined in terms of accuracy and precision along with the detection limit required, if applicable. This applies for process measurements, field instrument measurements, and analytical measurements.

For example, several process emission measurements may be required for evaluation of the technology. The DQO may be stated in terms of a regulatory limit. The investigator may wish to know not only if the limit is met but also the confidence associated with the reported measurement. A single reported number (e.g., 10 mg/m^3) has no associated error for the given measured quantity. The variability is determined by precision that is determined through repeated measurements. If total variability (sampling and instrument precision combined) is desired, then several measurements can be made in the field to determine means, standard deviations, and confidence limits of the reported results. If instrument variability needs to be separated from total error variability, replicate measurements of the same sample are required. These replicates establish a measure of precision.

Another measurement of error is accuracy. Accuracy is often determined by comparison to a known standard. Therefore, accuracy may be limited to indirect comparisons. For example, if particulate emissions are desired there may be no direct way to determine whether the collection method and subsequent determination are accurate. In this case, because a known standard may not exist, an approximate measure may be obtained through determination of the accuracy of the analytical instrumentation used to measure the final result. If a filter is used to collect the sample and subsequently weighed by a laboratory balance, then the accuracy of the balance used can be determined by a standard class S weight. In some cases, this may be the closest determination of error associated with accuracy and should be used as an approximation, as necessary.

For some measurements, a specific detection limit may be critical. For the above example, it may be important to know that the required regulatory limit can be measured. Instrument methods, therefore, need to be adjusted to measure the required detection limit. If it can't be measured by the chosen method, then modifications to the standard method are required or another method needs to be chosen.

QA objectives should be defined in terms of project requirements, and not in terms of the capabilities of the intended test methods. Test methods should be modified or chosen accordingly, such that QA requirements can be achieved. If the principal investigator is limited to a specific measurement method, then statistical confidence associated with specific project objectives may need modification.

The following is a step-wise procedure for determining appropriate QA objectives:

- Review the chosen objective and the confidence limit placed on that objective. To achieve these confidence limits (usually a statistical measure), a defined measurement error must be met. This defined measurement error is in terms of accuracy and precision as previously noted.
- Based on this error, choose the method that can meet this requirement for the specified measurement.
- Assuming that a standard method is to be used, determine that this method is indeed appropriate for determining the stated error in terms of accuracy and precision.
- If the total measurement error cannot meet specifications, then choose an alternate method or make appropriate alterations to the specified method.
- In some cases, it may be necessary to review the DQO if error specifications cannot be achieved. Perhaps a different DQO will be required because no measurement method exists that can meet the precision and accuracy requirements.
- Precision and accuracy are usually based on actual measurements and not on instrument specifications as stated by the manufacturer. For example, a flow meter may have specifications for precision of $\pm 1\%$ as determined by replicate measurements. It may also have a similar accuracy specification. Upon calibration, however, one may discover that accuracy is only within 5% compared to a known standard and the precision based on replicate measurements is only within 3%. These differences must be accounted for when determining stated DQOs. Accuracy in this case can probably be adjusted so that the reported value is calibrated accordingly. However, precision is a true measure of variability and is usually random error. Therefore, it is important to calibrate the chosen measuring device to ensure measured error is within specifications.

4.2 Completeness

Completeness is defined as the number of measurements judged valid compared to the number of measurements needed to achieve a specified level of confidence in decisionmaking. If it was estimated that eight measurements were needed to determine a 95% confidence level and if dust levels generated by the process meet regulatory requirements of less than $x \text{ mg/m}^3$, then additional measurements should be taken to ensure that this objective can be achieved. This is to ensure 10 valid data points will be available when evaluating the data. For example, it may be best to take 10 measurements assuming that at least eight valid measurements will exist for evaluation. Data can often be lost due to instrument or operator error. For this example, the completeness requirement would therefore be 80%.

4.3 Qualitative QA Objectives: Comparability and Representativeness

Comparability is the degree to which one data set can be compared to another. Comparability is achieved by the use of consistent methods and by traceability of standards to a

reliable source. Comparable data can only be obtained by quantitative measurements.

Representativeness is the degree to which a sample or group of samples is indicative of the population being studied. An environmental sample is representative for a parameter of interest when the average value obtained from a group of such samples tends toward the true value of that parameter in the actual environment, as the number of representative samples is increased, i.e., the value for that sample is within the Gaussian distribution of all samples. Representativeness is normally achieved by collecting a sufficiently large number of unbiased samples. This is usually defined by the statistical confidence needed for the stated objective(s) and the variability of the collected samples.

The QA objectives should demonstrate that adequate comparability and representativeness will be achieved.

4.4 What If QA Objectives Are Not Met?

A discussion of the impact of not meeting one or more of the previously specified QA objectives should be included. Will the project be a complete loss? Will some, but not all, of the project goals still be realized? Will the statistical confidence level be reduced? Answers to these types of questions help provide a perspective on how critical these objectives really are. If a QA objective will not be met during the course of a project, the project parameters, measurement methods, or even the overall project objectives need to be reevaluated or changed.

5. SAMPLING PROCEDURES AND DATA COLLECTION

This section must describe a plan for sampling that is responsive to those project objectives stated in Section 1. This section explains the overall sampling strategy and the specific sampling procedures that will be employed. It should also include a description of how field data will be documented.

5.1 CERCLA

Describe how the data collected will address the nine criteria identified in CERCLA. These criteria are important for the evaluation of alternatives. The nine criteria, with a short description of each, can be found in Table A-2. For the Landfill Stabilization Focus Area (LSFA), Criteria 2, 8, and 9 need not be addressed.

The information in Table A-2 is only a brief description of the criteria. For example, Implementability (Criterion 6) includes:

1. Technical feasibility, including technical difficulties and unknowns associated with the construction and operation, reliability, ease of undertaking additional remedial actions, and ability to monitor effectiveness.
2. Administrative feasibility, including activities needed to coordinate with other offices and agencies and the ability and time required to obtain any necessary approvals and permits from other agencies.
3. Availability of services and materials, including the availability of adequate offsite treatment, storage capacity, and disposal capacity and services; availability of necessary equipment and specialists, and provisions to ensure any necessary additional resources; availability of services and materials; and availability of prospective technologies.

5.2 Collection and Sampling Techniques

Describe the data collection and sampling techniques or guidelines used to collect test data. Identify the procedures to conduct the test, documents to be prepared detailing the procedures (i.e., detailed operating procedures), and applicable prerequisites, including regulatory documentation.

- Prepare a list of analytes, sample volumes to be collected, and the amount of sample required for analysis, if necessary.
- Describe splitting or compositing procedures used for samples. For example, compositing procedures may be used to reduce the number of samples needed for analysis.
- Describe sampling equipment that will be used and how this equipment will be calibrated. This includes field monitoring equipment.

Table A-2. Nine CERCLA criteria.

1. **Overall Protection of Human Health and the Environment.** This assessment is to determine whether the alternative can adequately protect human health and the environment, in both the short and long term, from unacceptable risks.
 2. **Compliance with ARARs.** This assessment is to determine whether the alternative can attain applicable or relevant and appropriate requirements (ARARs) under Federal, state, or facility (site) environmental laws or provide grounds for invoking a waiver.
 3. **Long-Term Effectiveness and Permanence.** The long-term effectiveness and permanence of each alternative is assessed, along with the degree of certainty that the alternative will prove successful.
 4. **Reduction of Toxicity, Mobility, and Volume Through Treatment.** This assessment is the degree to which the alternative employs recycling or treatment that reduces toxicity, mobility, or volume, including how treatment is used to address the principal threats posed by the site.
 5. **Short-Term Effectiveness.** This assessment is for the short-term impacts of alternatives.
 6. **Implementability.** This assessment is to identify the ease or difficulty of implementing the alternative.
 7. **Cost.** The assessment is to identify capital costs, including both direct and indirect, annual operation and maintenance (O&M) costs, and net present value of capital and O&M costs.
 8. **State Acceptance.** This is an assessment of state concerns and may not be completed until comments on the remedial investigation/feasibility study are received. Concerns may be discussed, however.
 9. **Community Acceptance.** This assessment includes determining which alternatives interested persons in the community support, have reservations about, or oppose. Additional information for each criterion can be found in EPA-540/G-89/004, Chapter 6, or Federal Register/Vol. 55, No. 46/Thursday, March 8, 1990/Rules and Regulations, pp. 8,849 and 8,850.
-

- Include a description of decontamination procedures as necessary.
- Include an explanation and an example of sample identification to use in the field.

5.3 Test Methods

Describe the test method(s) and uncertainties associated with these test methods identified at this point of the project, as scheduled. Uncertainties to be described include items like schedules and availability of sampling equipment, sampling rates, and application times, etc.

5.4 Sampling and Analysis Requirements

Sampling and analysis requirements may be identified in a separate sampling and analysis plan or included in the test plan/operating procedures. If the sampling and analysis plan is a subset of a larger demonstration, then reference the applicable sampling and analysis plan.

5.5 Sample Custody

Occasionally, samples are lost or compromised before or during sampling and analysis. Sample custody allows detection of such problems if they occur and minimizes such occurrences by assigning responsibility for all stages of sample handling. Sample custody is maintained when the samples are in a secure area or are in the view of, or under the control of, a particular individual. Records are maintained so that a sample history can be reconstructed later, if necessary. This section describes how sample custody will be maintained and recorded.

- Include names of sample custodians.
- Show examples of forms to be used for maintaining custody when samples are transported.
- Describe archiving of samples and shipping documents, including sample custody forms.

6. DOCUMENT CONTROL

6.1 Data

Describe how the data will be documented, controlled, and stored. If a database management system is used, include the frequency and method of data backup.

6.2 Test Plan

Explain how changes to the test plan will be documented, reviewed, and approved. This can be as simple as referring to an applicable procedure or configuration management plan that explains changes to documents (see Section 9, "Quality Assurance").

6.3 Other Documents

Describe document control methods for other documents used to perform the test, such as operating procedures, log books, etc. (see Section 9, "Quality Assurance").

7. ANALYTICAL METHODS, CALIBRATION PROCEDURES, AND LABORATORY QUALITY CONTROL

Describe the analytical method(s) to be used (if samples are to be analyzed). This should be a brief description. Reference standard methods or include copies of all other methods along with calibration procedures and appropriate laboratory quality control.

Describe and/or reference all specific quality control methods to be followed. Examples of items to be considered are replicates, spiked samples, split samples, control charts, blanks, and surrogate samples.

8. DATA REDUCTION, VALIDATION, AND VERIFICATION

The following data reduction, validation, and verification information is required:

- Describe the data reduction scheme, including all equations used to calculate the measured parameter and reporting units.
- Define how the collected data are validated.
- State who is responsible for validating the data.
- List and define the acceptance criteria. Explain how the acceptance criteria were derived.

9. QUALITY ASSURANCE

Quality assurance will help ensure that the data collected for an LSFA demonstration meet specifications of the DQOs and are meaningful. The following QA information is required:

- State the quality level, inspection requirements, and applicable quality program to which the test will adhere. Address how nonconformance will be resolved.
- Describe the routine procedures used to assess precision, accuracy, and completeness of measured data. These procedures should include the equations to calculate precision, accuracy, and completeness.

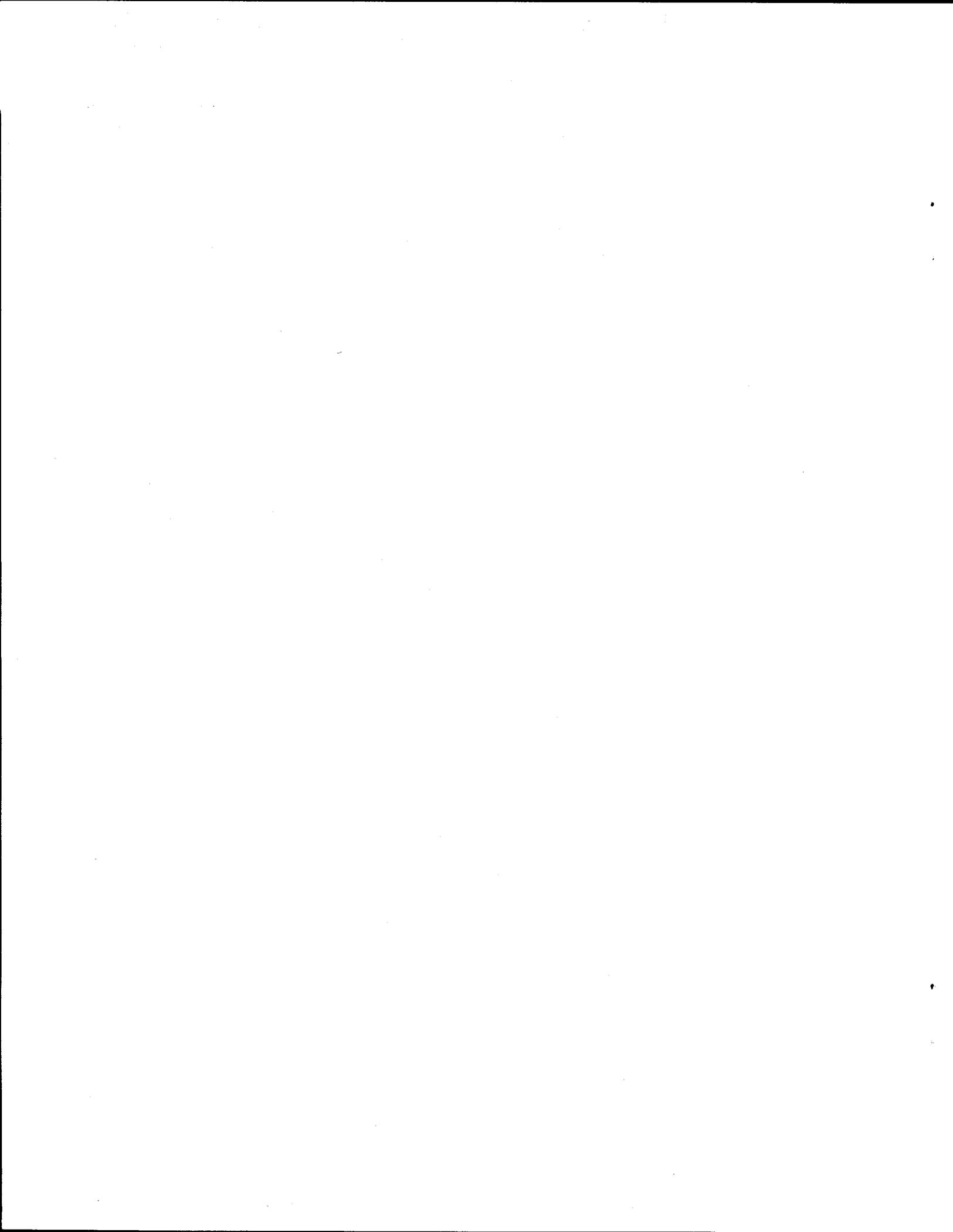
NOTE: Precision is the ability to repeat the data, accuracy is a measure of how close the data are to actual, and completeness is how much of the required data was obtained. Precision and accuracy are often identified as a percent and are determined by an equation/formula. For example, relative standard deviation can be used to assess precision for several replicate samples, or relative percent difference can be used to assess the precision for duplicate samples. Accuracy can be expressed as a percent of error from the true value. The definition for completeness is included in Section 3.

- Change control and corrective action.
 1. Describe how changes in equipment, procedures, documents, etc., will be completed and who can authorize the changes. Deciding who can authorize specific corrective action prior to the demonstration is important. A plan prevents both unauthorized changes and wasted time during a demonstration if it is determined that certain decisions can be made by field personnel. Describe the internal audit/surveillance system to verify that the proposed procedures are followed.
 2. Anticipate corrective action procedures that may be implemented. If certain changes in measurement equipment or process operations are a possibility, it is best to plan for these potential changes in advance.

10. REFERENCES

This section contains all references used in the test planning. Proper credit should be given to vendors and collaborating agencies.

Appendix B
Test Logistics and Deployment Guidance



Appendix B

Test Logistics and Deployment Guidance

1. Equipment and Instruments

- List the equipment and instruments required to run the test and who will furnish the equipment and instruments.
- Define range, accuracy, and tolerance requirements for the equipment and instruments.
- Describe the calibration or standardization procedure(s) for the equipment and instruments and include the frequency that the procedures will be performed.
- Define a preventive maintenance schedule for the equipment and instruments to minimize downtime. Include a list of critical spare parts.
- Describe any sensitivity to the environment (temperature extremes, dust, etc.).

2. Supplies, Utilities, and Facilities

Address the required supplies, utilities, and facility support to perform the test. This should include items such as forklifts, end loaders, storage buildings, office space, soil, boxes, fuel, telephones, electrical power, lights, craft support, etc. Include who will provide them.

3. Health and Safety

Identify the health and safety measures to ensure the health and safety of the operators and the surrounding community. This may be as simple as referencing the Regulatory Compliance Plan.

4. Residuals Management

Describe the type of waste by-products generated by the demonstration. Include the volume generated, physical state, and how the by-products will be disposed of.

5. Procedures

This appendix should include the detailed operating procedures, standard operating procedures, and any other required documents.