Nondestructive Evaluation: A Review of NDE Performance Demonstrations—NDE Round Robin Report

1016969

Nondestructive Evaluation: A Review of NDE Performance Demonstrations—NDE Round Robin Report

1016969

Technical Update, June 2008

EPRI Project Manager

C. Miller

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

Electric Power Research Institute (EPRI)

CITATIONS

This document was prepared by

Electric Power Research Institute (EPRI)
Nondestructive Evaluation (NDE) Program
1300 W.T. Harris Blvd.
Charlotte, NC 28262

Principal Investigator
C. Miller

This document describes research sponsored by EPRI.

This publication is a corporate document that should be cited in the literature in the following manner:

Nondestructive Evaluation: A Review of NDE Performance Demonstrations—NDE Round Robin Report. EPRI, Palo Alto, CA: 2008. 1016969.

PRODUCT DESCRIPTION

History has shown that re-qualification examinations for intergranular stress corrosion cracking (IGSCC) ultrasonic (UT) examiners have pass rates averaging 57%. An alternative to stringent re-qualification requirements is to provide annual practice or training as part of a performance-based qualification approach. This report summarizes the past performance of experienced candidates who have been certified to different written practices and programs and indicate their capabilities as measured by undertaking hands-on practical qualification testing.

Results and Findings

Various performance demonstrations conducted worldwide were evaluated to determine the ability of certified nondestructive evaluation (NDE) personnel to pass performance-based practical tests and demonstrations. These tests were conducted throughout the world for a variety of industries. In general, the pass rates for these other blind tests have been comparable to the nuclear domestic experience.

Challenges and Objectives

Utility personnel responsible for in-service inspections have become concerned with the decrease in quantity, quality, and knowledge of the NDE work force. This is evidenced, for example, by the numerous attempts necessary to successfully pass the UT performance tests required by the nuclear industry since May 2000. The pass rate obtained by experienced certified personnel is rarely above 50% on the first attempt at a demonstration test that simulates the examinations conducted in the field. Some of the test mockups with flaws were even removed from service, so there can be no question about their applicability to field inspections.

Applications, Value, and Use

There is certainly room for improvement in the area of effectively qualifying NDE personnel. Adopting performance-based concepts to the training, experience, and testing of NDE personnel should improve the reliability of examinations in the field.

EPRI Perspective

The effect of low pass rates has a large economic impact on the costs to maintain the qualification of NDE personnel. Each attempt by a candidate to pass a demonstration test incurs tuition, labor, and usually per-diem costs. Because candidates who do not pass the qualification tests may not be hired for nuclear work, the individual loses income and the employer is not compensated for preparation training that was provided.

Approach

Additional supporting evidence was obtained during this study that substantiates the concern that certified NDE personnel may not have adequate training and experience to perform reliable inspections for applications in a variety of industries. A literature review was conducted to evaluate the ability of certified personnel to pass performance-based practical tests and demonstrations. These tests were conducted throughout the world for a variety of industries.

Keywords
NDE
Performance demonstration
Human factors
Round robin testing

ACKNOWLEDGEMENTS

EPRI would like to thank Mike Turnbow from the Tennessee Valley Authority for his assistance in providing documentation for some of the studies included in the report and for having the foresight to initiate this study to support progress towards performance based nondestructive examination personnel certifications.

ABSTRACT

This report describes a study performed by the Electric Power Research Institute (EPRI) Nondestructive Evaluation (NDE) Program to assist the Tennessee Valley Authority in providing a basis for performance-based NDE personnel certification. History has shown that re-qualification examinations for intergranular stress corrosion cracking (IGSCC) ultrasonic (UT) examiners have pass rates averaging 57%. The U.S. Nuclear Regulatory Commission (NRC) staff has questioned whether examiners have maintained their proficiency over the three-year re-qualification period and may recommend more frequent and more stringent testing requirements. An alternative to stringent re-qualification requirements is to provide annual practice or training as part of a performance-based qualification approach. This report will summarize the past performance of experienced candidates that have been certified to different written practices and programs and indicate their capabilities as measured by undertaking hands-on practical qualification testing. Each chapter describes a study and includes an excerpt from the study's documentation.

CONTENTS

1 **INTRODUCTION 1-1**

2 **NONDESTRUCTIVE EVALUATION CAPABILITIES DATA BOOK (THIRD EDITION)........ 2-1**

2.1 NDE Process and Process Variance 2-2
2.2 Results of Liquid Penetrant Testing by Three Operators 2-4
2.3 Magnetic Particle Inspection 2-5
2.4 Results of Magnetic Particle Testing by Three Operators 2-6
2.5 X-Radiographic Inspection 2-7
2.6 Results of X-Radiography Testing by Three Operators 2-8
2.7 Eddy Current Inspection 2-9
2.8 Results of Eddy Current Testing by Three Operators 2-11
2.9 Ultrasonic Inspection 2-12
2.10 Results of Ultrasonic Testing by Three Operators 2-13
2.11 Lessons Learned in Application of POD Methods 2-14
2.12 Summary 2-14

3 **RELIABILITY OF NONDESTRUCTIVE INSPECTION (NDI) OF AIRCRAFT COMPONENTS 3-1**

3.1 Background 3-1
3.2 Scope 3-1
3.3 Facility, Equipment and Processing Effects on Performance 3-2
3.4 Fluorescent Penetrant Inspection 3-2
3.5 Human Factors Effects on Inspection Performance 3-3
3.5.1 General 3-3
3.5.2 Human Factors Observations from AFLC Assessments 3-3
3.5.3 Human Factors Summary 3-7

4 **AMERICAN PETROLEUM INSTITUTE (API) QUALIFICATION OF ULTRASONIC EXAMINERS CERTIFICATION PROGRAM 4-1**

4.1 Purpose 4-1
4.2 Testing Protocol 4-1
4.3 Specimen Presentation 4-1
4.4 Test Set and Test Specimen Design 4-1
4.5 Potential Flaw Mechanisms 4-2
4.5.1 Specimen Criteria 4-3
4.6 Grading Criteria 4-3
4.6.1 Equipment Requirements 4-3
4.7 Results from Previous Examination Sessions 4-4

5 **PANI II SEMINAR REPORT 5-1**

6 **WHITE PAPER REPORT 1 .... 6-1**
6.1 Introduction 6-1
6.2 Industry Efforts to Improve Performance 6-1
6.3 Conclusions 6-2

7 **WHITE PAPER REPORT 2 .... 7-1**
7.1 Introduction 7-1
7.2 The Essentials of Efficient Skills Learning 7-1
7.3 New Technologies and Resources for Skills Development 7-2
7.4 A Structured Approach to NDE Skills Development 7-2
7.5 Implementation 7-3
7.6 Conclusion 7-3

8 **PERFORMANCE DEMONSTRATION IN THE USA (PISC III) 8-1**
8.1 Objective 8-1
8.2 Procedures 8-1
8.3 Results 8-1
8.4 Piping 8-2
8.5 RPV 8-4
8.6 Conclusions 8-5

9 **PROGRAM FOR THE ASSESSMENT OF NDT IN INDUSTRY (PANI I) 9-1**
9.1 Objective 9-1
9.2 Procedures 9-1
9.3 Results 9-2
9.4 Manufactured Test Pieces 9-2
9.5 Discussion 9-4
9.6 Conclusions 9-4

10 **PERFORMANCE DEMONSTRATION – 25 YEARS OF PROGRESS 10-1**
10.1 Abstract 10-1
10.2 Introduction 10-1
10.3 Solutions 10-2
10.4 Performance Demonstration 10-2
10.5 Progress 10-2
10.6 Discussion 10-5

11 **REACTOR PRESSURE VESSEL INSPECTION RELIABILITY BASED ON PERFORMANCE DEMONSTRATIONS 11-1**
11.1 Background 11-1
11.2 History 11-1
11.3 Objectives 11-2
11.4 Effectiveness of Examinations 11-2
11.5 Passed-Plus-Failed Candidates 11-3

11.6 Probability of Detection 11-5
11.7 Data Interpretation...11-6

***12* COMPARISON OF PERFORMANCE TEST RESULTS WITH FIELD PERFORMANCE MEASURES...12-1**

12.1 Background 12-1
12.2 Conclusions and Recommendations..12-2

***13* REFERENCES 13-1**

LIST OF FIGURES

Figure 2-1 Operator A Test Results 2-4
Figure 2-2 Operator B Test Results 2-4
Figure 2-3 Operator C Test Results 2-4
Figure 2-4 Operator A Test results 2-6
Figure 2-5 Operator B Test Results 2-6
Figure 2-6 Operator C Test Results 2-6
Figure 2-7 Operator A Test Results 2-8
Figure 2-8 Operator B Test Results 2-8
Figure 2-9 Operator C Test Results 2-8
Figure 2-10 Operator A Test Results 2-11
Figure 2-11 Operator B Test Results 2-11
Figure 2-12 Operator C Test Results 2-11
Figure 2-13 Operator A Test Results 2-13
Figure 2-14 Operator B Test Results 2-13
Figure 2-15 Operator C Test Results 2-13
Figure 3-1 POD Curve for a Group IV Penetrant Inspection Sequence that was Performed by an Inspector who had Spent 75% of His Time on the Penetrant Line 3-5
Figure 3-2 POD Curve for a Group IV Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned to Penetrant Inspection for Twenty Days during the Year 3-5
Figure 3-3 POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned Full Time to Penetrant Inspection 3-5
Figure 3-4 POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned on a Fill-In Basis Only 3-6
Figure 3-5 POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned Full Time to Penetrant Inspection 3-6
Figure 3-6 POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who was Assigned to Penetrant Inspection on a Fill-In Basis Only (Weekends and Holidays) 3-6
Figure 4-1 Test Specimen Weld Layouts 4-2
Figure 4-2 Test Specimens 4-2
Figure 6-1 Comparison of PDI Candidate POD and Three Round Robin Exercises Material is Austenitic Piping 11 to 25 mm in Thickness 6-2
Figure 8-1 Manual Piping Pass Rates 8-2
Figure 8-2 Passing Rates by Attempt 8-3
Figure 8-3 Comparisons of PDI and NDEC IGSCC 8-3
Figure 8-4 Summary of Inside Surface Demonstrations 8-4
Figure 8-5 Summary of Outside Surface RPV Demonstrations 8-5
Figure 9-1 Operator Performance in Detecting Defects with Ultrasonics in Manufactured Test Pieces ... 9-3
Figure 9-2 Detectability of Individual Defects in Manufactured Test Pieces 9-3
Figure 10-1 Probability of Detection for ID Examinations, for Flaws at the Clad-to-Base-Metal Interface, Based on PDI Demonstration Results 10-3
Figure 10-2 RPV Sizing Error Distributions for PDI and PISC II Advanced Techniques 10-4
Figure 10-3 Comparison of PDI Candidate POD and Three Round Robin Exercises. Material is Austenitic Piping 11 to 25 mm (0.43 to 0.98 inches) in Thickness 10-5
Figure 11-1 POD for Passed and Passed-Plus-Failed Candidates for Automated Examinations from the Inside Surface 11-3
Figure 11-2 POD With 95% Confidence Bounds for Passed Candidates 11-4

LIST OF TABLES

Table 2-1 Dominant Sources of Variance in NDE Procedure Application 2-2
Table 2-2 Liquid Penetrant Use Variables 2-3
Table 2-3 Magnetic Particle Inspection Use Variables 2-5
Table 2-4 X-Radiographic Inspection Use Variables 2-7
Table 2-5 Eddy Current Manual (Hand) Scan Inspection Use Variables 2-9
Table 2-6 Eddy Current Automated Scan Inspection Use Variables 2-10
Table 2-7 Ultrasonic Inspection Use Variables 2-12
Table 3-1 Automatic Penetrant Processing Black Light Intensity and Detection Levels 3-3
Table 3-2 Test Set 2, Penetrant Inspection Results Showing Inspector Performance Variations 3-4
Table 3-3 Relationship of Recent Experience and Inspection Performance 3-4
Table 4-1 Potential Flaw Mechanisms 4-3
Table 9-1 Operator Performance in Detecting Defects with Ultrasonics in Manufactured Test Pieces 9-2
Table 11-1 Number of Detection Measurements 11-6
Table 12-1 Comparison Between Average Percentages of ODSCC Indications Detected at Steam Generator Tube-Support Locations from Performance Tests (Test) and Field Performance Measures (Field) 12-2

1 INTRODUCTION

History has shown that re-qualification examinations for intergranular stress corrosion cracking (IGSCC) ultrasonic (UT) examiners have pass rates averaging 57%. The U.S. Nuclear Regulatory Commission (NRC) staff has questioned whether examiners have maintained their proficiency over the three-year re-qualification period and may recommend more frequent and more stringent testing requirements. An alternative to stringent re-qualification requirements is to provide annual practice or training as part of a performance-based qualification approach. This report will summarize the past performance of experienced candidates that have been certified to different written practices and programs and indicate their capabilities as measured by undertaking hands-on practical qualification testing.

Personnel are commonly qualified for technical tasks by completing a specified amount of training and time on the job (on-the-job training). However, exposure to training and job experience does not necessarily lead to job proficiency. For example, research involving 32 candidates (two groups) for qualification in the Electric Power Research Institute (EPRI) Nondestructive Evaluation (NDE) Center Performance Demonstration Initiative (PDI) program revealed no positive relationship at all between the number of years of NDE, ultrasonic, or piping examination experience and performance on ultrasonic examination demonstration tests. Results such as these are not puzzling if one recognizes that unstructured training and job experience may or may not include the key factors that are essential for learning: measurement and feedback of performance results. It is difficult to measure performance and provide feedback on the job—actual flaws may be rarely encountered. Timely, accurate measurement and feedback may be difficult even in laboratory-type blind demonstration tests but if demonstration tests of similar difficulty and conditions are provided under a training setting then the skills should be learned by the candidates. This performance based training will likely provide the desired result of improving the personnel examination skills.

Utility personnel responsible for in-service inspections have become concerned with the decrease in quantity, quality, and knowledge of the NDE workforce. This is evidenced, for example, by the numerous attempts necessary to successfully pass the UT performance tests required by the nuclear industry since May 2000. Additional supporting evidence has been obtained in the course of this study to substantiate the concern that certified personnel may not have adequate training and the experience to perform reliable inspections for particular applications in a variety of industries.

As part of this study, a literature review was conducted to evaluate the performance of certified personnel in their ability to pass performance-based practical tests and demonstrations. The remainder of this report consists of a chapter for each study that was conducted. The chapter will include a brief synopsis of the study including the NDE method tested, excerpts from the study, author's opinions, test protocol, mockup design, and performance of the personnel. These tests were conducted worldwide for a variety of industries. In general the pass rates for these other blind tests have been comparable to the nuclear domestic experience. That is, the pass rates obtained by experienced certified personnel is very rarely above 50% on their first attempt at passing a demonstration test that simulates the examinations that they conduct in the field. Some of the test mockups with flaws were even removed from service so there can be no question about their applicability to field inspections.

These pass rates have a large impact on the costs to maintain the qualification of NDE personnel. Each attempt by a candidate to pass a demonstration test incurs “tuition”, labor, and usually per-diem costs. Candidates that don’t pass the qualification tests may not be hired for nuclear work so the individual loses income and the employer is not compensated for preparation training that was provided. There is room for improvement in the area of effectively qualifying NDE personnel. Adopting performance-based concepts to training, experience and testing of NDE personnel in the future should improve the reliability of examinations in the field.

2

NONDESTRUCTIVE EVALUATION CAPABILITIES DATA BOOK (THIRD EDITION)

This Nondestructive Evaluation (NDE) Capabilities Data Book [1] is intended to be a baseline for engineering analyses in the form of a condensed reference to previously demonstrated NDE capabilities for the aircraft industry. It is intended to be a companion to damage tolerance and safe-life analysis tools that are integral to quantitative design practices and “fracture control” and safe-life of engineering hardware.

Damage tolerance and safe-life analyses provide a quantified basis for structural integrity in the form of a “critical crack size” or “assumed crack size” required for life assessment. The most desirable assumption would be “no cracks” and many engineering documents specify “no cracks” as a requirement for acceptance. Experience, economics, and materials behavior do not, however, support a “no cracks” criterion as a basis for safe-life design.

A firm requirement to quantify and demonstrate the capabilities of NDE procedures was imposed by the development and application of fracture mechanics in damage tolerance design practices. Fracture mechanics analysis produced a single valued “critical crack size” as a basis for damage tolerance analysis. Detection of cracks below the critical crack size is required for the implementation of “fracture critical” and safe-life analysis. NDE involves multiple process variables and does not produce a single valued result. Variances in NDE process capabilities were addressed using the tools that were developed for establishing materials properties design values in the form of the “probability of detection” (POD) as a function of crack / flaw size. Since a wide range of NDE methods and procedures are used in “fracture control” of engineering hardware and systems, a large volume of POD data have been generated to validate the capabilities of specific NDE procedures in a multitude of applications.

POD data are specific to particular NDE procedures and applications and do not readily fit the “Handbook” paradigm. Condensation of POD data is therefore presented in the form of this Data Book and users are cautioned that the data are valuable for purposes of reference and general understanding of NDE capabilities, but specific NDE capabilities validation data and disciplines must be generated by the user to support critical hardware design and use.

In summary, the purpose of this Data Book is to provide a condensation of quantified capabilities data that have been developed and documented in previous applications. Its intended use is a single point reference to results that have been obtained in state-of-the-art NDE applications and therefore provide a link to prior art. The modular format of the Data Book anticipates data additions that may include additional NDE processes and procedures and various procedure applications.

2.1 NDE Process and Process Variance

The diverse nature of different NDE processes results in different sources of variance and resultant impact on detection output capabilities. For example, a manually applied liquid penetrant process is dominated by the skill of the operator in process application and interpretation. An automated eddy current process is dominated by calibration, instrument, and procedure variances. It is important to recognize the source of variance in each NDE process and to take the nature of the variance and process control into account in applying margins to the NDE processes. Table 2-1 shows typical dominant sources of variance for the following NDE processes:

- Liquid penetrant inspection
- Magnetic particle inspection
- Radiographic inspection (X-ray and gamma ray)
- Electromagnetic inspection
- Ultrasonic inspection

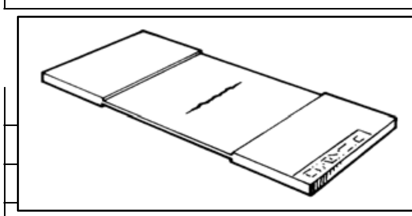
Table 2-1
Dominant Sources of Variance in NDE Procedure Application

	Materials	Equipment	Procedure	Calibration	Criteria	Human Factors
Liquid Penetrant	X		X			X
Magnetic Particle	X	X	X			X
X-ray	X	X	X			X
Manual Eddy Current		X	X	X	X	X
Automatic Eddy Current		X	X	X	X	
Manual Ultrasonic		X	X	X	X	X
Automatic Ultrasonic		X	X	X	X	

NDE methods and procedures are selected using a variety of practical implementation criteria. The lowest cost method that produces the required result is usually the method of choice. Part of the cost consideration must be the cost of qualifying/validating the capabilities of the method and the cost of maintaining such qualification. The data included herein are intended to be an aid to method/procedure selection and implementation.

Table 2-2
Liquid Penetrant Use Variables

C`1000(1L	DATA SET DESCRIPTION (PT - 01 (1))
METHOD:	Fluorescent Penetrant
TEST OBJECT TYPE:	Flat Plate - 3.5 inches by 16 inches, cracks on both sides
NDE PROCEDURE:	Fluorescent Penetrant Manual - URESCO P149, Solvent Removable, Spray Developer
ARTIFACT TYPE:	Fatigue Cracks - R< 0. 70 (Shaped EDM starter notch initiation, growth in bending and tension / tension
ARTIFACT SHAPE:	ASPECT RATIO - 0.1 TO O.5 (a/2c) --- DEPTH TO THICKNESS - 0.2 TO 0.5 (a/t)
ARTIFACT VERIFICATION:	Destructive analysis and measurement
MATERIAL:	2219 Aluminum T-87
TEST OBJECT THICKNESS:	0.060 and 0.225 inch nominal
TEST OBJECT CONDITION:	-01,"As Machined", -02,"After Etch", -03,"After Proof"
SURFACE FINISH:	125 and 32 RMS - representative of good machining practices
APPLICATION:	Manual Inspection / Manual Recording
DATA SET IDENTIFIER:	PTAAA01-A,B,C; PTAAA02-A,B,C; PTAAA03-A,B,C
TYPE OF DATA:	Hit / Miss with estimated crack lengths
TEST OPPORTUNITIES:	311 Cracks
DETECTED:	PTAAA01-A= 187, B= 173, C= 233
FALSE CALLS:	Not reported
REFERENCE:	NASA CR-2369 Rummel, Ward D., Paul H. Todd Jr., Sandor A. Frecska, and Richard A. Rathke, <u>The Detection of Fatigue Cracks by Nondestructive Testing Methods</u>, February 1974.
DATE:	November 1971 - June 1973
WORK SPONSOR:	W.L. Castner, NASA Lyndon B. Johnson Space Center
PERFORMING ORGANIZATION:	Martin Marietta Aerospace, Denver, Colorado
NOTES:	This program was performed in support of the National Aeronautics Administration (NASA) Space Shuttle design and is the first known publication of nondestructive evaluation data in a continuous function probability of detection (POD).
	Flaws were induced in 105 panels (both sides). Thirteen blank panels were included for a total of 118 panels
	The original data analysis was in the form of a moving average plot. Data have been reanalyzed and plotted here by the maximum likelihood / log logistic method.
	A parallel program was conducted by the General Dynamics Corp, San Diego, CA.; test panels were exchanged and inspections repeated by both organizations.



90% POD	"AS MACHINED"	"AFTER ETCH"	"AFTER PROOF"
	A= 0.390 in.	A= 0.089 in.	A= 0.017 in.
	B= 0.435 in.	B= 0.077 in.	B= 0.059 in.
	C= 0.313 in.	C= 0.106 in.	C= 0.077 in.

2.2 Results of Liquid Penetrant Testing by Three Operators

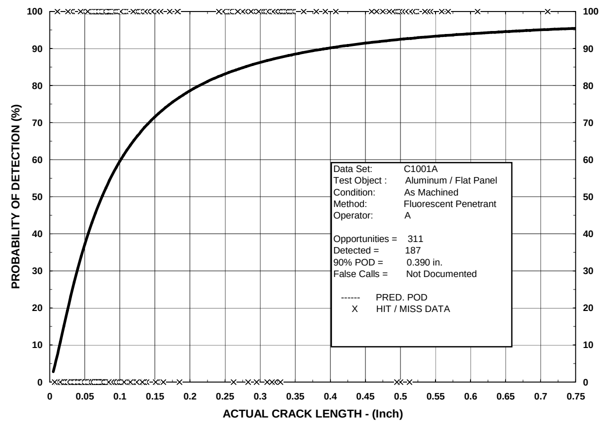


Figure 2-1
Operator A Test Results

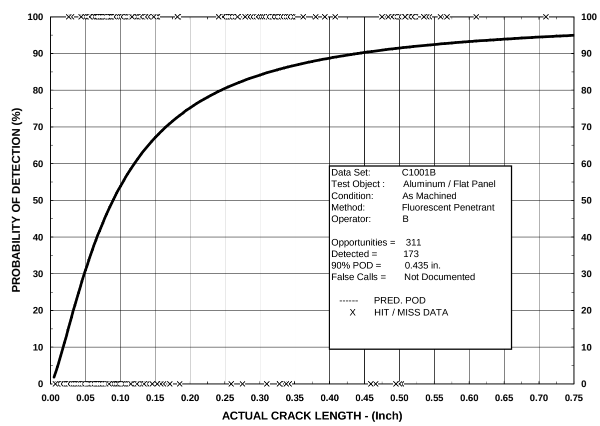


Figure 2-2
Operator B Test Results

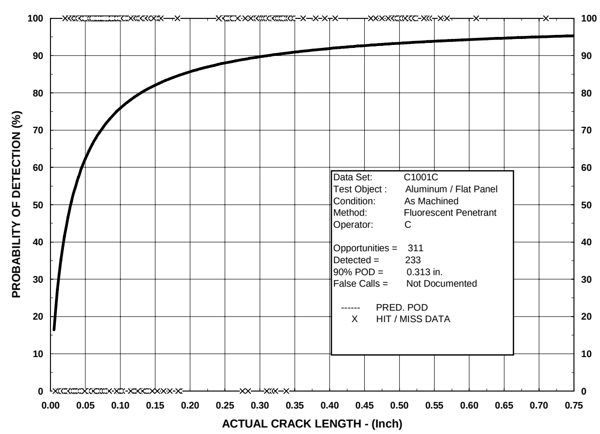
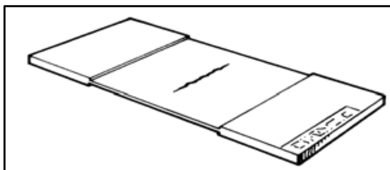


Figure 2-3
Operator C Test Results

2.3 Magnetic Particle Inspection

Table 2-3
Magnetic Particle Inspection Use Variables

B1000(2D	DATA SET DESCRIPTION
METHOD:	Magnetic Particle Inspection by CRACK DEPTH
TEST OBJECT TYPE:	Flat Plate - 6 inches by 16 inches, cracks on both sides
NDE PROCEDURE:	Eddy Current - Contact probe; 3 MHz, Meter read-out
ARTIFACT TYPE:	Fatigue Cracks - R< 0. 70 (Shaped EDM starter notch initiation, growth in bending and tension / tension
ARTIFACT SHAPE:	ASPECT RATIO - 0.1 TO O.5 (a/2c) --- DEPTH TO THICKNESS - 0.2 TO 0.5 (a/t)
ARTIFACT VERIFICATION:	Destructive analysis and measurement
MATERIAL:	Steel - 4340
TEST OBJECT THICKNESS:	0.060 and 0.250 inch nominal
TEST OBJECT CONDITION:	-01,"As Machined", -02,"After Etch", -03,B1"After Proof"
SURFACE FINISH:	125 RMS - representative of good machining practices
APPLICATION:	Manual Processing / Manual Inspection (Wet Horizontal / Uresco 228 fluorescent particles)
DATA SET IDENTIFIER:	B1001.A,B,C; B1003,A,B,C
TYPE OF DATA:	Hit / Miss with estimated crack lengths
TEST OPPORTUNITIES:	120 Cracks - As Machined; 132 Cracks After Etch and Proof Load
DETECTED:	MT-01D - A = 97, B= 102, C= 104
FALSE CALLS:	Not reported (<5%)
REFERENCE:	NASA CR-151098, Rummel, Ward D., Richard A. Rathke, Paul H. Todd Jr., Thomas L. Tedrow, and Steve J. Mullen, **Detection of Tightly Closed Flaws by Nondestructive Testing Methods in Steel and Titanium**, November, 1976.
DATE:	July 1975 - September 1976
WORK SPONSOR:	W.L. Castner, NASA Lyndon B. Johnson Space Center
PERFORMING ORGANIZATION:	Martin Marietta Aerospace, Denver, Colorado
NOTES:	This program was performed in support of the National Aeronautics Administration (NASA) Space Shuttle design and was used as a basis for design / acceptance criteria. Flaws were induced in 45 panels (both sides). Fifteen (15) blank panels were included for a total of 60 panels The original data analysis was in the form of a moving average plot. Data have been reanalyzed and plotted here by the maximum likelihood / log logistic method.



CRACK DEPTH

90% POD	"AS MACHINED"	"AFTER PROOF"
	A= Not Achieved	A= Not Achieved
	B= Not Achieved	B= Not Achieved
	C= Not Achieved	C= 0.043in. (1.105mm)

2.4 Results of Magnetic Particle Testing by Three Operators

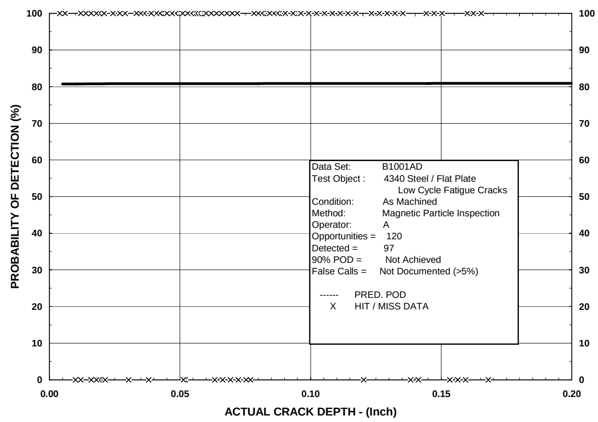


Figure 2-4
Operator A Test results

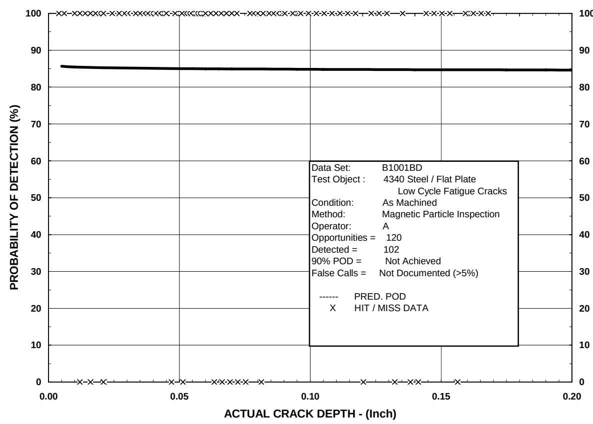


Figure 2-5
Operator B Test Results

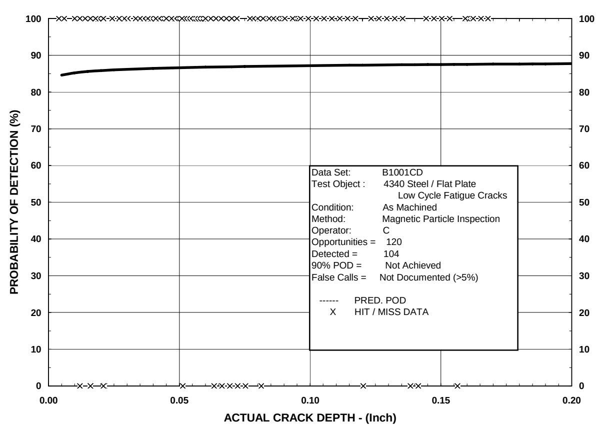
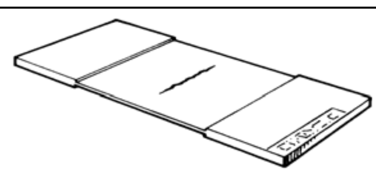


Figure 2-6
Operator C Test Results

2.5 X-Radiographic Inspection

Table 2-4
X-Radiographic Inspection Use Variables

F4060(2L	**DATA SET DESCRIPTION**
METHOD:	X-Radiographic Inspection by CRACK LENGTH - 0.060" panel thickness
TEST OBJECT TYPE:	Flat Plate - 6 inches by 16 inches, cracks on both sides
NDE PROCEDURE:	X-radiographic inspection; Kodak Type M Film; Automatic Processing; Exposure optimized to test specimens
ARTIFACT TYPE:	Fatigue Cracks - R< 0. 70 (Shaped EDM starter notch initiation, growth in bending and tension / tension
ARTIFACT SHAPE:	ASPECT RATIO - 0.1 TO O.5 (a/2c) --- DEPTH TO THICKNESS - 0.2 TO 0.5 (a/t)
ARTIFACT VERIFICATION:	Destructive analysis and measurement
MATERIAL:	Steel - 4340
TEST OBJECT THICKNESS:	0.060 inch nominal
TEST OBJECT CONDITION:	F40601,"As Machined"; F40603,"After Etch and Proof Loading"
SURFACE FINISH:	125 RMS - representative of good machining practices
APPLICATION:	Manual Processing / Manual Inspection (Wet Horizontal / Uresco 228 fluorescent particles)
DATA SET IDENTIFIER:	F40601.A,B,C; F40603,A,B,C
TYPE OF DATA:	Hit / Miss with estimated crack lengths
TEST OPPORTUNITIES:	80 Cracks -Variation in the number inspected during each sequence
DETECTED:	F40601 - A = 22/62, B= 22/62, C= 23/62
FALSE CALLS:	Not reported (<5%)
REFERENCE:	NASA CR-151098, Rummel, Ward D., Richard A. Rathke, Paul H. Todd Jr., Thomas L. Tedrow, and Steve J. Mullen, Detection of Tightly Closed Flaws by Nondestructive Testing Methods in Steel and Titanium, November, 1976.
DATE:	July 1975 - September 1976
WORK SPONSOR:	W.L. Castner, NASA Lyndon B. Johnson Space Center
PERFORMING ORGANIZATION:	Martin Marietta Aerospace, Denver, Colorado
NOTES:	This program was performed in support of the National Aeronautics Administration (NASA) Space Shuttle design and was used as a basis for design / acceptance criteria.
	Flaws were induced in 45 panels (both sides). Fifteen (15) blank panels were included for a total of 60 panels
	The original data analysis was in the form of a moving average plot. Data have been reanalyzed and plotted here by the maximum likelihood / log logistic method.
	CRACK LENGTH
	90% POD "AS MACHINED" "AFTER PROOF"
	A= Not Achieved A= 0.263in. (6.67mm)
	B= Not Achieved B= 0.486in. (12.35mm)
	C= Not Achieved) C= 0.263in. (6.67mm)

2.6 Results of X-Radiography Testing by Three Operators

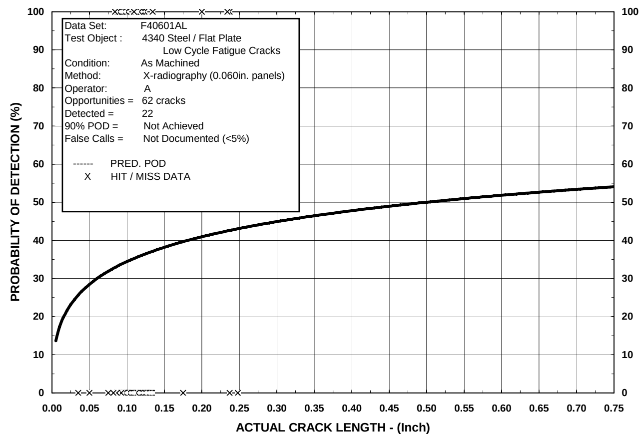


Figure 2-7
Operator A Test Results

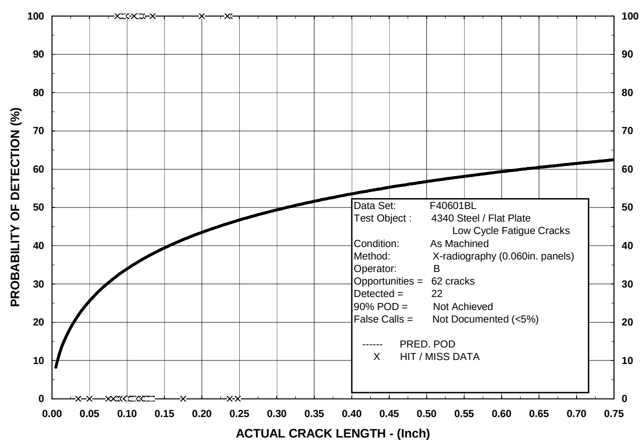


Figure 2-8
Operator B Test Results

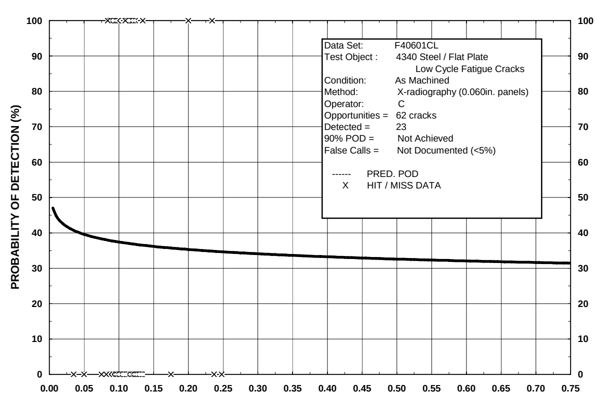


Figure 2-9
Operator C Test Results

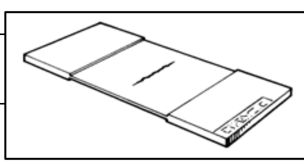
2.7 Eddy Current Inspection

Table 2-5
Eddy Current Manual (Hand) Scan Inspection Use Variables

MANUAL (HAND) SCAN

COST OF INSPECTION	LOW
COST OF EQUIPMENT	MODERATE
OPERATOR SKILL REQUIREMENTS	HIGH
PROCESS CONTROL REQUIREMENTS	MODERATE
PROCESS VARIANCE / MARGIN REQUIREMENTS	MODERATE

Table 2-6
Eddy Current Automated Scan Inspection Use Variables

A1000(1L	DATA SET DESCRIPTION (ET - 01 (1) CRACK LENGTH)
METHOD:	Eddy Current
TEST OBJECT TYPE:	Flat Plate - 3.5 inches by 16 inches, cracks on both sides
NDE PROCEDURE:	Eddy Current - Contact Probe 100 kHz, Meter Readout
ARTIFACT TYPE:	Fatigue Cracks - R< 0. 70 (Shaped EDM starter notch initiation, growth in bending and tension / tension
ARTIFACT SHAPE:	ASPECT RATIO - 0.1 TO O.5 (a/2c) --- DEPTH TO THICKNESS - 0.2 TO 0.5 (a/t)
ARTIFACT VERIFICATION:	Destructive analysis and measurement
MATERIAL:	2219 Aluminum T-87
TEST OBJECT THICKNESS:	0.060 and 0.225 inch nominal
TEST OBJECT CONDITION:	-01,"As Machined", -02,"After Etch", -03,B1"After Proof"
SURFACE FINISH:	125 and 32 RMS - representative of good machining practices
APPLICATION:	Hand Scanning - Manual Readout
DATA SET IDENTIFIER:	ETAAA01-A,B,C; ETAAA02-A,B,C; ETAAA03-A,B,C
TYPE OF DATA:	Hit / Miss with estimated crack lengths
TEST OPPORTUNITIES:	311 Cracks
DETECTED:	ETAAA01-A= 208, B= 224, C= 205; 02-A= 228, B= 273, C= 243; 03-A= 264, B= 268, C= 266
FALSE CALLS:	Not reported
REFERENCE:	NASA CR-2369 Rummel, Ward D., Paul H. Todd Jr., Sandor A. Frecska, and Richard A. Rathke, The Detection of Fatigue Cracks by Nondestructive Testing Methods, February 1974.
DATE:	November 1971 - June 1973
WORK SPONSOR:	W.L. Castner, NASA Lyndon B. Johnson Space Center
PERFORMING ORGANIZATION:	Martin Marietta Aerospace, Denver, Colorado
NOTES:	This program was performed in support of the National Aeronautics Administration (NASA) Space Shuttle design and is the first known publication of nondestructive evaluation data in a continuous function probability of detection (POD).
	Flaws were induced in 105 panels (both sides). Thirteen blank panels were included for a total of 118 panels
	The original data analysis was in the form of a moving average plot. Data have been reanalyzed and plotted here by the maximum likelihood / log logistic method.
	A parallel program was conducted by the General Dynamics Corp, San Diego, CA.; test panels were exchanged and inspections repeated by both organizations.
	90% POD Length - "AS MACHINED" "AFTER ETCH" "AFTER PROOF"
	A= 0.196 in. A= 0.198 in. A= 0.052 in.
	B= 0.184 in. B= 0.071 in. B= 0.037 in.
	C= 0.295 in. C= 0.270 in. C= 0.0871 in.

2.8 Results of Eddy Current Testing by Three Operators

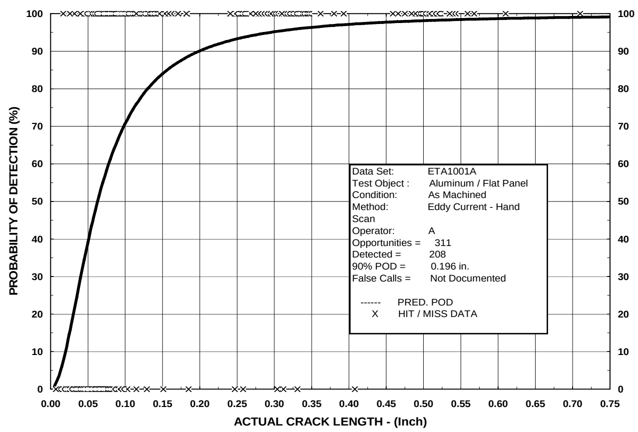


Figure 2-10
Operator A Test Results

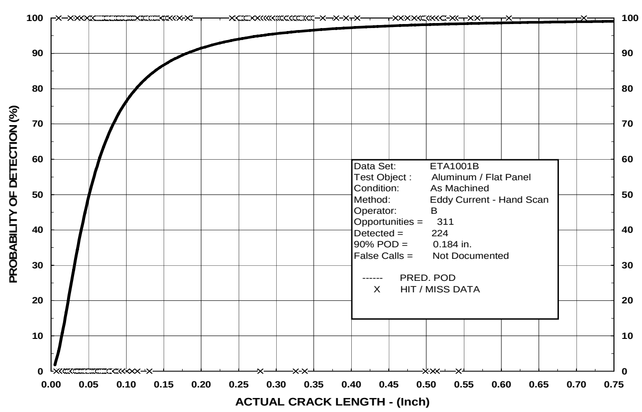


Figure 2-11
Operator B Test Results

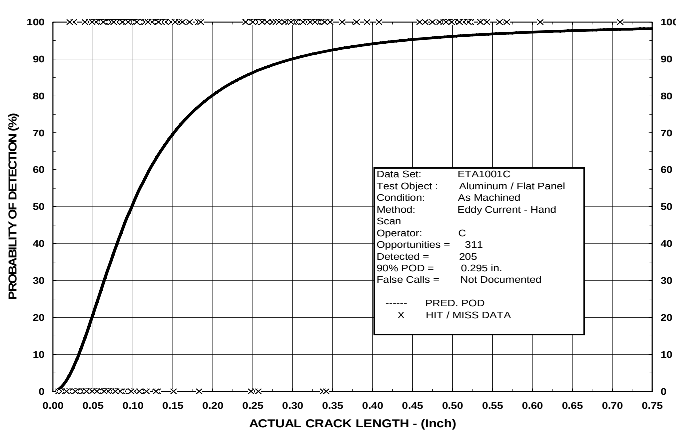
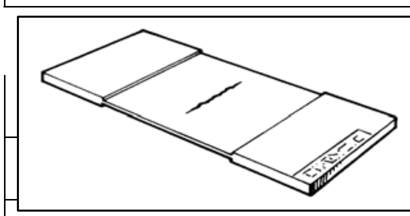


Figure 2-12
Operator C Test Results

2.9 Ultrasonic Inspection

Table 2-7
Ultrasonic Inspection Use Variables

D1000(1D	DATA SET DESCRIPTION (UT - 01 (1) CRACK DEPTH)
METHOD:	Ultrasonic Surface Wave
TEST OBJECT TYPE:	Flat Plate - 3.5 inches by 16 inches, cracks on both sides
NDE PROCEDURE:	Ultrasonic Surface Wave - Immersion at 10MHz
ARTIFACT TYPE:	Fatigue Cracks - R< 0. 70 (Shaped EDM starter notch initiation, growth in bending and tension / tension
ARTIFACT SHAPE:	ASPECT RATIO - 0.1 TO O.5 (a/2c) --- DEPTH TO THICKNESS - 0.2 TO 0.5 (a/t)
ARTIFACT VERIFICATION:	Destructive analysis and measurement
MATERIAL:	2219 Aluminum T-87
TEST OBJECT THICKNESS:	0.060 and 0.225 inch nominal
TEST OBJECT CONDITION:	-01,"As Machined", -02,"After Etch", -03,"After Proof"
SURFACE FINISH:	125 and 32 RMS - representative of good machining practices
APPLICATION:	Immersion - C-scan recording
DATA SET IDENTIFIER:	UTAAA01D-A,B,C; UTAAA02D-A,B,C; UTAAA03D-A,B,C
TYPE OF DATA:	Hit / Miss with estimated crack lengths
TEST OPPORTUNITIES:	311 / 306 Cracks (5 cracks lost in proof test)
DETECTED:	UTAAA01D-A= 258, B= 261, C= 254;
FALSE CALLS:	Not reported
REFERENCE:	NASA CR-2369 Rummel, Ward D., Paul H. Todd Jr., Sandor A. Frecska, and Richard A. Rathke, **The Detection of Fatigue Cracks by Nondestructive Testing Methods**, February 1974.
DATE:	November 1971 - June 1973
WORK SPONSOR:	W.L. Castner, NASA Lyndon B. Johnson Space Center
PERFORMING ORGANIZATION:	Martin Marietta Aerospace, Denver, Colorado
NOTES:	This program was performed in support of the National Aeronautics Administration (NASA) Space Shuttle design and is the first known publication of nondestructive evaluation data in a continuous function probability of detection (POD).
	Flaws were induced in 105 panels (both sides). Thirteen blank panels were included for a total of 118 panels
	The original data analysis was in the form of a moving average plot. Data have been reanalyzed and plotted here by the maximum likelihood / log logistic method.
	A parallel program was conducted by the General Dynamics Corp, San Diego, CA.; test panels were exchanged and inspections repeated by both organizations.
	90% POD Depth - "AS MACHINED" "AFTER ETCH" "AFTER PROOF"
	A= 0.025 in. A= 0.056 in. A= 0.014 in.
	B= 0.025 in. B= 0.016 in. B= 0.015 in.
	C= 0.024 in. C= 0.025 in. C= 0.031 in.

2.10 Results of Ultrasonic Testing by Three Operators

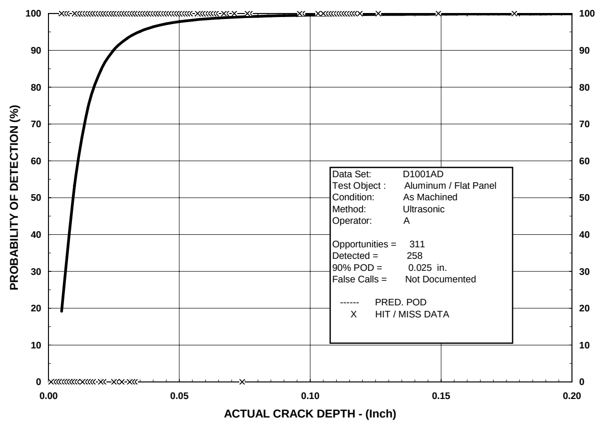


Figure 2-13
Operator A Test Results

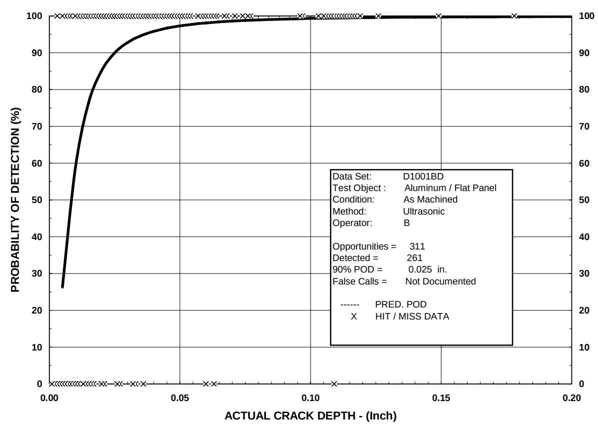


Figure 2-14
Operator B Test Results

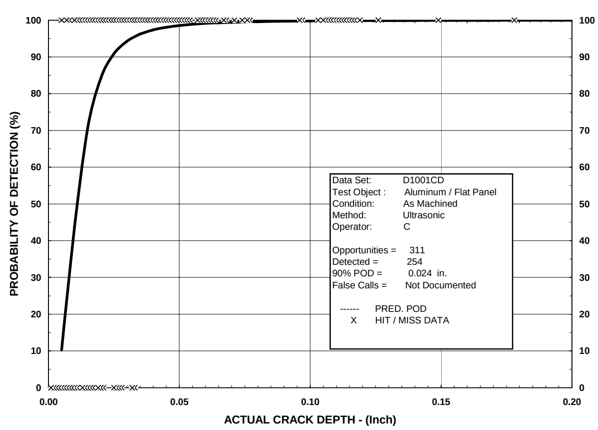


Figure 2-15
Operator C Test Results

2.11 Lessons Learned in Application of POD Methods

Lessons learned by application of POD methods to NDE procedure characterization include:

- NDE capabilities are rarely at the levels assumed by deterministic process management.
- Human factors in NDE are important, but are most often not the weak link in NDE procedure application.
- NDE capabilities are rarely at the level assumed by use of a “calibration” artifact of a fixed size. (Note: calibration artifacts (“standards”) are extremely important in establishing a baseline set-point and a linearity response check for instrumented NDE methods.)
- Calibration artifacts are a primary source of variance in the reproducibility and repeatability of NDE procedures that are conducted at different locations.
- “Standard” NDE procedures should be periodically assessed to assure that the procedure is producing the desired result. (Note: The 29/29 point estimate test is a rapid method of providing assurance of basic and continuing NDE procedure performance.)
- Not all NDE procedures require characterization/quantification. The use of the procedure may reduce the need for characterization.
- The POD method is integral to meeting critical damage tolerance requirements.

2.12 Summary

NDE processes have evolved as essential tools in the production, acceptance, life-cycle operations, and maintenance of modern engineering materials, components, structures, and systems. In many applications, a traditional "best effort" or workmanship confidence in NDE procedure application is adequate to provide the required level of discrimination and confidence in "fitness for purpose" of the test object. As demands for NDE procedures have increased, new methods and increased precision in application have been realized, and reliance on NDE procedure application has increased to produce increasing efficiencies in engineering designs and to add confidence in continuing fitness for purpose of aging systems. While the basic principles of NDE processes are simple in concept as described in previous sections of this chapter, precision and confidence level in application require in-depth understanding and control of NDE process parameters. Quantification of NDE procedure capabilities is a complex process and a measure of the maturity of NDE technology. Excellence in NDE procedure development, qualification, application and management are essential to capable and reliable NDE process application. This Data Book is dedicated to the goal of excellence in NDE engineering and NDE process application.

3
RELIABILITY OF NONDESTRUCTIVE INSPECTION (NDI) OF AIRCRAFT COMPONENTS

3.1 Background

This report documents and concludes an initial task to assess the reliability of nondestructive inspection (NDI) of aircraft engine components in Air Force engine overhaul facilities [2]. Air Force Logistics Center (AFLC) excerpts from the original report have been included in this EPRI report. The work constitutes one of the most comprehensive efforts yet undertaken to assess, quantify, and document nondestructive inspection performance capabilities and reliability in a production line environment. The primary purpose, goal, and value of the program was to provide an independent baseline assessment of AFLC production line nondestructive inspection operations. The baseline assessment was accomplished by sampling a number of specimens for each operation and using a statistically significant sample size to quantify the effectiveness of each operation. Rigor in analysis was achieved for those inspection operations that were under control at the time of assessment. Several process control improvements were identified and recommendations were made to improve overall process control, inspection capability, and reliability. Some process control improvement recommendations were made by direct observations in the production line and some were made as a result of off-line analyses. Improvements resulting from direct observations were made, where possible, and additional sampling was performed to quantify the effect of the improvement on overall inspection performance.

Secondary goals of the program were to describe and document the methodology, observations, data gathering, and data analyses and to recommend improvements such that performance improvements may be quantified at a future date. In addition, observations and “lessons learned” were to be documented and reported such that future assessment programs may benefit from the logic, methodology, and baseline data established during this program. Reported direct observations have resulted in the implementation of some performance improvements in materials assessment, equipment and maintenance improvements, process improvements, and overall operational improvements.

3.2 Scope

The scope of the program was limited to initial tear-down inspection operations for components in the gas path. Inspection methods assessed included:

- Fluorescent Penetrant Inspection (FPI)
- Eddy Current Inspection (EC)
- Ultrasonic Inspection Testing (UT)
- Magnetic Particle Testing (MT)

Since the majority of the inspection operations in Air Force engine maintenance facilities are performed using fluorescent penetrant processes, special attention was given to fluorescent penetrant inspection in the NDI reliability assessment program. X-radiographic inspection was not included since it is performed off-line by laboratory personnel.

3.3 Facility, Equipment and Processing Effects on Performance

Penetrant inspection processing parameters have a significant impact on the brightness of flaw indications which can be identified by the inspector. Some processing variations were due to operator skills; some were due to poor nondestructive engineering of the inspections; some were due to facility supplies and support, and some were due to processing equipment characteristics. If the processing parameters specified to the inspector have not been validated as being capable of producing the required inspections, the inspector may not have a chance of performing at the levels expected. Facility, equipment, and processing parameters that were noted during assessment of nondestructive inspection capabilities and reliability are discussed below.

3.4 Fluorescent Penetrant Inspection

Fluorescent penetrant inspection equipment, processes, and processing parameters vary considerably throughout the AFLC facilities. The most notable facility discrepancy was that of not supplying qualified penetrant inspection materials and not supplying adequate material to replace discrepant materials. Some of the observed processing parameters variations were the result of limitations in the facilities and equipment. Other variations were the result of the lack of adequate processing instructions and to human factors variables.

- Inspection Materials—Qualified penetrant inspection materials were not available. This was attributed to procurement policies and to the failure to perform acceptance tests on receipt of materials.
- Penetrant Dwell Times—Penetrant dwell times observed at the AFLC facilities were generally between 20 and 30 minutes. There were a few cases where extended dwell times in excess of 90 minutes were noted. In one of these cases, a water wash penetrant was used which dried on the part. This inspection resulted in 232 false calls in 240 opportunities. There were no cases noted where insufficient dwell time resulted in poor inspection performance.
- Emulsification Time—The emulsification times noted at the AFLC facilities were widely variable ranging from less than 30 seconds to well over 6 minutes. Variation in emulsification time in hand processes were due to a lack of processing instructions and to human factors variation. Processing parameters on the automated lines did not match the parameters for the materials procured.
- Dry Developer Application and Exposure Times—Some dry developer processing stations were not functioning properly at the time of the assessment. This was due in part to equipment design and part to failure to maintain the equipment. The dry developer exposure time was also found to have a significant effect on the overall inspection performance and was, in most cases, inadequate for high performance.
- Wet Developer Application—Wet developer materials were discrepant at both facilities. In addition, the level of fluorescence at some of the wet developer stations was high and the material was in need of replacement.

- Black Light Intensities—Black light intensities were found to be below specified levels at numerous inspection stations and the assessment results showed this had an effect on the inspection performance. A direct correlation between black light intensity and performance was found on Test Set 2 using automatic penetrant processing. All inspections were performed using the same materials and identical processing times. Table 3-1 lists some detection percentages and black light intensities showing the relationship between black light intensity and inspection performance.

Table 3-1
Automatic Penetrant Processing Black Light Intensity and Detection Levels

Test Set	Black Light Intensity (microwatts/sq. cm.)	Detection Percentages	Opp.	F.C.
2	590	18%	276	10
2	728	26%	276	0
2	830	36%	276	8
2	954	50%	276	0

3.5 Human Factors Effects on Inspection Performance

3.5.1 General

The overall performance level for a nondestructive inspection operation is dependent on the adequacy of the NDI engineering and acceptance criteria definition; the NDI materials; equipment; processes; and human skills applied to the operation. Previous studies have cited human factors variations as the primary factor in NDI performance. Although human factors variables were observed, other inspection variables dominated the AFLC facilities assessments. In short, unless the NDI engineering, materials, equipment, and processes are under control, the human operator at the end of the line will have a very difficult time performing satisfactorily.

3.5.2 Human Factors Observations from AFLC Assessments

The performance of individual inspectors was variable. To show the amount of variation, Test Set 2 (titanium panels simulating anti-rotation slots) was inspected by seven different operators on one inspection line using identical materials and processing parameters. There are 276 opportunities in this test set. The detection results from inspections were performed using poor performing developer and extended emulsification times, the results are not a measure of inspector capability, but do however show the variation in capabilities between individual inspectors (and materials).

Table 3-2
Test Set 2, Penetrant Inspection Results Showing Inspector Performance Variations

Test Set	Inspector Number	Detection Percentages	Number of False Calls
2	1	37.90%	15
	2	35.50%	9
	3	50.00%	0
	4	30.60%	12
	5	25.80%	0
	6	17.70%	10
	7	24.20%	10

In reviewing the inspector profiles obtained from each inspector prior to each inspection, there was a consistent and significant correlation between performance and the inspectors recent inspection experience. Those inspectors who performed penetrant inspections on a daily basis consistently out-performed those inspectors who performed penetrant inspections on an irregular or fill-in basis only on the same test set. Figures 3-1 through 3-6 are POD curves from three pairs of inspections that clearly show this relationship. Each pair of inspections was performed on the same line using like materials and processing times. Table 3-3 lists the detection percentages and number of false calls for each of these inspections with the proportion of time spent per week on penetrant inspections by the inspector.

Table 3-3
Relationship of Recent Experience and Inspection Performance

Test Set	Inspector No. (Figure)	Time Assigned to to Penetrant Insp.	Detection %	Number of False Calls	OPP.
3	**8**	30 hrs/week	76%	7	232
3	**9**	20 days/year	60%	3	232
4	**10**	40 hrs/week	50%	2	240
4	**11**	Fill in, weekends/Holidays	30%	7	240
5	**12**	40 hrs/week	81%	14	180
5	**13**	Fill in, weekends	59%	3	180

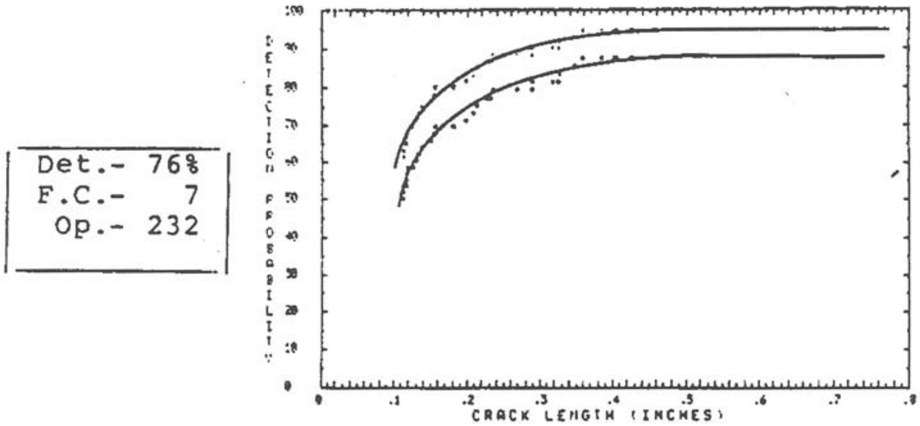


Figure 3-1
POD Curve for a Group IV Penetrant Inspection Sequence that was Performed by an Inspector who had Spent 75% of His Time on the Penetrant Line

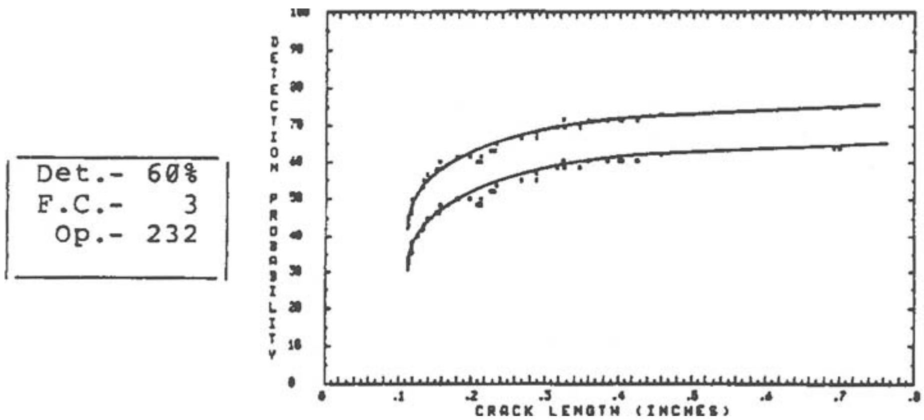


Figure 3-2
POD Curve for a Group IV Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned to Penetrant Inspection for Twenty Days during the Year

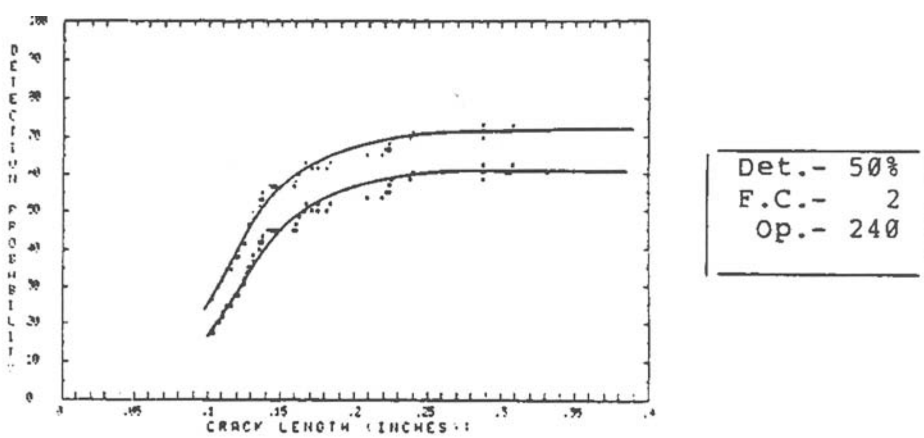


Figure 3-3
POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned Full Time to Penetrant Inspection

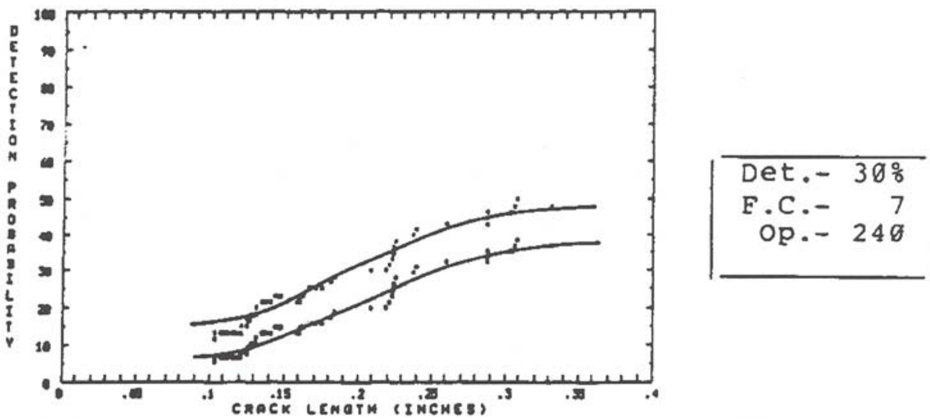


Figure 3-4
POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned on a Fill-In Basis Only

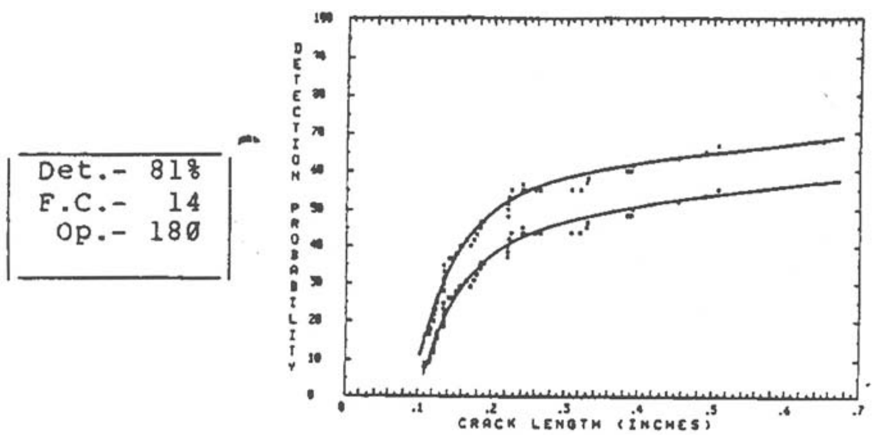


Figure 3-5
POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who had been Assigned Full Time to Penetrant Inspection

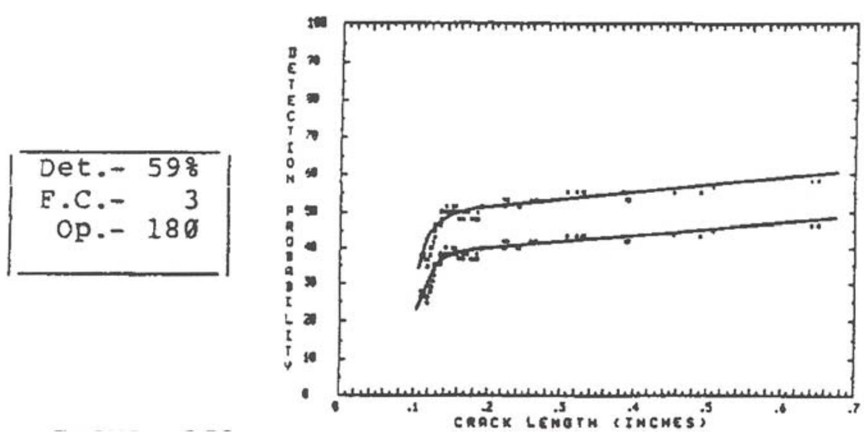


Figure 3-6
POD Curve for a Group IVa Penetrant Inspection Sequence that was Performed by an Inspector who was Assigned to Penetrant Inspection on a Fill-In Basis Only (Weekends and Holidays)

There was no clear correlation found between length and amount of training/experience and performance. As a whole though, those inspection areas, where an active on-the-job (OJT) program was being utilized, maintained a higher level of proficiency than did those areas without such a program.

3.5.3 Human Factors Summary

The human operator is usually reliable in performing a wide variety of tasks. The human operator must, however, operate within the boundary conditions of his capabilities and within the boundary conditions set by the physical limits of the task to be performed. NDE tasks are complex and require knowledge, skill, experience, and dexterity for optimum performance. When detection and discrimination are not attained by an NDI process, the most frequent cause reported is that of operator error. Operator errors, due to factors discussed n this paper, can and do occur. Proper attention must be given to those NDI processes involving human operators to assure that the tools, working conditions, and environment are commensurate with the task being performed.

4 AMERICAN PETROLEUM INSTITUTE (API) QUALIFICATION OF ULTRASONIC EXAMINERS CERTIFICATION PROGRAM

4.1 Purpose

The information in this study is intended to provide an outline of the API Qualification of Ultrasonic Examiners Certification Program [3]. This is intended to provide a brief overview regarding test administration and candidate preparation.

4.2 Testing Protocol

Each candidate was given a unique test set consisting of four qualification test specimens. Candidates must work independently and are not allowed to discuss specimen or examination information during or after the demonstration. In general, there was no single sample time limit established, however, if a sample requires sharing between two candidates, sample time limit provisions were established.

4.3 Specimen Presentation

Flaw location and "True" specimen identification were concealed to maintain a "blind test". Test specimens were given a unique identifier or alias. Test specimens are divided into grading units. Each grading unit was considered as either flawed or unflawed. The number of flawed and/or unflawed grading units and specific grading unit length was not made available to the candidates. There was no disclosure of particular specimen results or candidate viewing of unmasked specimens during or after the performance demonstration.

4.4 Test Set and Test Specimen Design

Each candidate was supplied a list defining the test specimens that make up their test set. Each test set was comprised of the following samples (one of each):

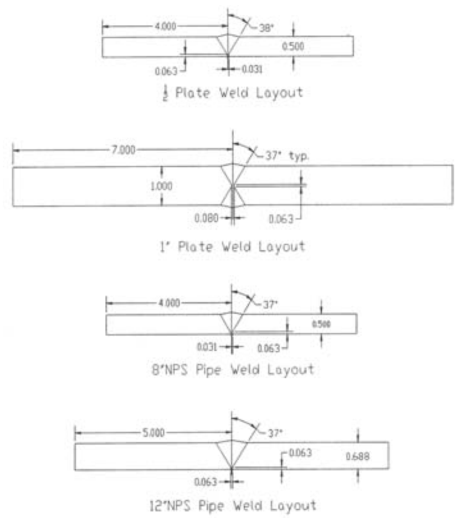


Figure 4-1
Test Specimen Weld Layouts

Figure 4-2
Test Specimens

4.5 Potential Flaw Mechanisms

The following table identifies the potential flaw mechanisms that may be included in each test specimen. The number of flaws in each test specimen may vary for each test set. Test specimens may be unflawed along the entire length.

Table 4-1
Potential Flaw Mechanisms

	½" Plate	1" Plate	8" Pipe Weld	12" Pipe Weld
Inside surface connected crack (ID Crack)	X	X	X	X
Outside surface connected crack (OD Crack)	X	X	X	X
Embedded Center Line Cracking		X		
Lack of root penetration (LOP)	X		X	X
Lack of side wall fusion (LOF)	X	X	X	X
Porosity	X	X	X	X
Slag inclusion	X	X	X	X

4.5.1 Specimen Criteria

- Specimens will not contain counterbore geometry.
- Specimens may contain ID or OD mismatch.
- Specimen weld crowns and root geometries will be in the as-welded condition and may be offset from sample centerline.

4.6 Grading Criteria

Candidate performance was evaluated in the following four areas:

- Detection - The detection portion of the test is applied to initially evaluate a candidate's data report. If the candidate does not detect an intended flaw, no further evaluation is required. The candidate will be required to detect approximately of 80% of the flaws in the test set. Sufficient data must be provided in order for the monitor to determine if the candidate actually detected the flaw.
- Flaw Characterization - Once the intended flaw is determined to be a "detection", the candidate's ability to characterize the flaw will be evaluated. Candidates will be provided with a list of potential flaw mechanisms for each sample type. Characterization criteria will be weighted heavily on the location of the reported flaw (surface connected or volumetric). The candidate must correctly characterize approximately 80% of detected flaws.
- Flaw Positioning - Reported flaws must also be positioned correctly with respect to the weld centerline. Cross sectional plotting of flaw indications on the indication data sheets will be required in order to determine the location of the flaw. The candidate must correctly position approximately 80% of the detected flaws.
- False Calls - A false call is defined as the reporting of a flaw (regardless of length) within a non-flawed grading unit. Candidates will not know the location of the unflawed grading units. The candidate must correctly evaluate approximately 80% of the unflawed grading units (20% false call rate) in order to be successful.

4.6.1 Equipment Requirements

- Candidate or candidate organizations are responsible for supplied all the equipment needed for each demonstration. Sharing of equipment will not be allowed during the demonstration.

4.7 Results from Previous Examination Sessions

- 295 Examinations
- 136 Passing
- 155 Failing
- 46.10% Pass Rate

5
PANI II SEMINAR REPORT

By John Thompson

A majority of this chapter includes excerpts from Mr. Thompson's report on this seminar and the opinions presented here are not the opinions of EPRI.

The seminar [4] took place on December 11, 2003, at Risley, and was attended by around 80 delegates. Overall, the seminar was well organized and informative. Delegates were provided with summaries of the presentations and the program for the day was as follows:

1. Opening remarks by John Whittle, Chairman of the Program for the Assessment of NDT in Industry (PANI) Management Committee
2. Introduction, by Harry Bainbridge, HSE
3. Results, by Bernard McGrath, Serco Assurance
4. Analysis, by Geoff Worral, Serco Assurance
5. Operator/vendor view, by Matt Clark, Belmont NDT
6. PANI Management Committee by Roger Lyon, RWE Innogy
7. Open discussion session
8. Closing remarks by John Whittle, Chairman of the PANI Management Committee

30 operators had been grouped according to the desire to determine:

- results using the "standard NDT procedure" based on EN 1714 (control group)
- results using an "improved (different) procedure"
- results following training (unspecified)

Each group of 10 operators had tested 6 specimens:

- Nozzle weld (P2101)
- Nozzle weld (P2102)
- Pipe, with 4 butt welds and a bend with erosion (P2301)
- 'T' fillet weld (P2021)
- 'J' weld (P2041
- Specimen from plant with real root crack (P2051)

If the trial had first tested 10 operators using the standard procedure, then the same 10 operators using the "improved" procedure, then the same 10 operators following meaningful training, the results would have been truly indicative of the effects of these "improvements".

The program demonstrated that overall detection rates for manual UT are of the order of 50%, spread between 30% and 80% detection. This is in line with all other UT round robin exercises that Thompson is familiar with.

According to Thompson, taking into account the Personnel Certification in Non-Destructive Testing (PCN) recertification experience, where the majority of candidates undergo some form of refresher training prior to examination, and even so, failure rates are ~ 30% to 40%. This should come as no surprise. Furthermore, the above rates are worse when one takes into consideration the fact that a significant number of recertification candidates hold multiple group certifications (a combination of plate, pipe, ‘T’, nozzle and node), and come way with less groups than were held prior to recertifying. These latter candidates are not amongst the 30 – 40% regarded as failing to recertify! The data available from PCN examinations is a potentially valuable resource.

No clear proposals emerged from discussion at the seminar. So, what are the conclusions and recommendations? Well, the program is not yet completed, and will take into consideration the feedback from industry. Nevertheless, in the view of Thompson, there is no perfect solution to the problem – 100% detection can never be a reality – and the only recourse is to “manage” and thus reduce the problem to an acceptable level. What is an acceptable level though?

Managing the problem will require additional measures, the cost of which would have to be borne by industry. These could include:

1. Supervised practice. Operators need to demonstrate to themselves, their employers, and to recipients of their UT services, that they are capable of detecting >95% of significant defects, and this could be achieved through a program of testing flawed specimens under supervision in a training environment, where failure to achieve the desired detection rate would result in refresher training. The question is: what should be the periodicity of such exercises? Six monthly? Twelve monthly?
2. Refresher training. Many UT operators hold multi group certification (a combination of plate, pipe, ‘T’, nozzle and node), and often do not carry out inspection of one or more of the certified groups for considerable periods of time. Operators could be required to log their ongoing experience and, when a particular group is not tested during, say, a 12 month period, some appropriate and relevant refresher training could become mandatory.
3. Repeat testing. It has been established that detection rates are substantially improved within a regime of repeat testing or cross checking. There are two main reasons for this: the first, obvious, reason must be simply that each operator will miss defects, but not necessarily the same defects; the second could be that, knowing a repeat inspection will or could be carried out, each operator is more vigilant because no one wants to be shown to be less than competent. Clearly, the cost involved is high, so this would only be adopted for critical applications where the consequence of failure is unbearable. It could also be implemented on a random and un-notified basis in order to “keep operators sharp”. Deficiencies noted during such “cross checks” would result in refresher training. There is nothing new in this!

There may be other more simple measures that could prove to be cost effective. For example, the PCN renewal requirements stipulate a regime of (at least) annual surveillance conducted internally by the employer. It would be interesting to know which of the 30 operators were working within such a regime, and to analyze their performance in the trial in comparison with those who were not working within such a regime!

What action can the British Institute of Non-Destructive Testing (BINDT)/PCN take? We shall need to address concerns with some form of action, but should await the publication of the formal report before deciding what, if anything, needs to be done at certification scheme level.

Once the PANI II report is published, we should arrange similar meetings to those following PANI I in order that we consult and garner views from as wide a cross section as possible.

6
WHITE PAPER REPORT 1

What is the Real Quality of ASNT Certified Personnel?
Jack Spanner Jr.

OBJECTIVE: the objective of this paper is to summarize Spanner's thoughts on the performance of NDE personnel certified to ASNT's SNT-TC-1A when participating in various performance demonstration programs in preparation for panel discussion at the ASNT 2002 Fall Conference. [5]

6.1 Introduction

A recent study [6] found that only about 24 percent of the current NDE workforce will still be available for NDE work by the year 2010. Utility personnel responsible for in-service inspections have become concerned with the decrease in quantity, quality, and knowledge of the NDE workforce. This is evidenced, for example, by the numerous attempts necessary to successfully pass the ultrasonic (UT) performance tests required by the nuclear industry since May 2000. Prior to that UT examiners were required to pass a demonstration program beginning in 1983 that included pipe sections containing IGSCC that had been removed from service. In general, the pass rate for IGSCC detection on the first attempt was between 19% and 34% and this increased to 27% and 70% after multiple attempts. This data is based on 27 classes with 182 candidates held from September 1985 until July 1987 [7]. Only 18% of the candidates passed the combined General, Specific and Practical tests on the first attempt. The IGSCC qualified personnel are still required to pass the IGSCC demonstration test **without the written tests** every three years and their pass rate is approximately 50% on the first attempt. This increases to approximately 75% after multiple attempts. Nearly all of the participants in these two programs are fully certified to ASNT Level II or Level III in the UT method!

Obviously, the effect of these pass rates have a large economic impact on the costs to maintain the qualification of NDE personnel. Each attempt by a candidate to pass a demonstration test incurs "tuition" costs, labor, and usually per-diem costs. Candidates that don't pass the qualification tests may not be hired for nuclear work so the individual loses income and the employer is not compensated for preparation training that was provided.

6.2 Industry Efforts to Improve Performance

Several ultrasonic round robins have been conducted since 1982 and there has been a steady improvement in performance. Figure 6-1 compares the POD for PDI piping candidates with those from three previous round robin exercises [8]. These include the Piping Inspection Round Robin (PIRR) 1981-1982 [9], the Mini Round Robin (MRR) 1986 [10], and the PISC-AST 1989-1990 [11]. The data presented for PDI represents the same pipe sizes and flaw types as were included in the other three studies. The PDI data include results from stainless steel piping with a wall thickness range of 11-25 mm. Flaws types in the PDI data include thermal fatigue and field removed IGSCC. The PDI data include passed candidates and passed-plus- failed candidates over the years 1994 to 2001. Manual examinations make up the large majority of the PDI data included in Figure 6-1. The nuclear industry has improved the reliability of their inspections by improving the training and using performance based procedures instead of prescribed procedures. The certification of these personnel has not changed significantly so that cannot be the reason for the improvements.

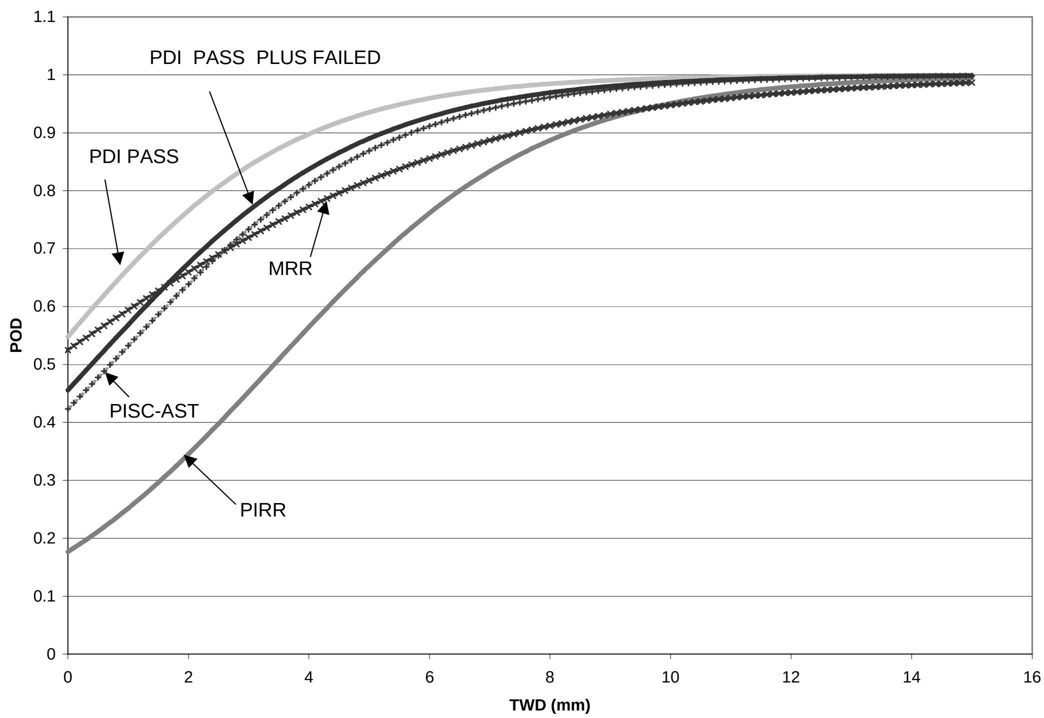


Figure 6-1
Comparison of PDI Candidate POD and Three Round Robin Exercises
Material is Austenitic Piping 11 to 25 mm in Thickness

6.3 Conclusions

It is reasonable to assume that other industries are experiencing similar results when using performance demonstrations to qualify personnel, procedures, and equipment. For example only 30% of the certified personnel that attempted to pass the American Petroleum Institute's ultrasonic flaw detection proficiency test were successful [12]. Requirements for the certification of NDE personnel should be revised based on these experiences. A round robin to validate the effectiveness of the revisions should then be conducted.

7
WHITE PAPER REPORT 2

Proposed Performance Based Training and Qualification

Jack Spanner Jr.

OBJECTIVE: to describe a new system developed by Spanner for NDE personnel training and qualification that is performance based to replace the current prescriptive system in preparation for participation in a panel discussion at ASNT's 2003 spring conference. [13]

7.1 Introduction

While retirement will be a principal factor in future NDE workforce reduction, more than 60% of the current NDE workforce is expected to leave the nuclear power industry for reasons other than retirement. Two of the reasons frequently cited include problems that add frustration and difficulty to the work itself; excessive qualification requirements and inappropriate rules and regulations. Another problem faced by the industry in filling this gap is the long lead times required to develop qualified NDE personnel, commonly estimated to be at least three years under the current system of classroom and on-the-job training. One approach to meeting future NDE personnel requirements is to increase the efficiency and effectiveness of developing NDE personnel, and thus reduce the lead times required to develop the capabilities required. This paper describes how this might be done through the use of new technologies and the structuring of practice and job experience. This would not only alleviate difficulties associated with retaining the workforce but would also provide a positive effect on recruiting.

7.2 The Essentials of Efficient Skills Learning

Personnel are commonly qualified for technical tasks by completing a specified amount of training and time on the job (on-the-job training). However, exposure to training and job experience does not necessarily lead to job proficiency. Recent research involving 32 candidates (two groups) for qualification in the PDI program [14] revealed no positive relationship at all between the number of years of NDE, ultrasonic, or piping examination experience and ultrasonic-examination performance. This also was recently confirmed by a Swedish study [15]. Results such as these are not puzzling if one recognizes that unstructured training and job experience may or may not include the key factors that are essential for learning-<u>measurement and feedback of performance results</u>. Admittedly, it is difficult to measure performance and provide feedback on the job-actual flaws may be rarely encountered-and timely, accurate measurement and feedback may be difficult even in laboratory-type training settings. It is important, therefore, to look to new technologies as avenues for incorporating key learning factors into the skill development process, and to focus on these factors in structuring both training and on-the-job experience. Personnel learn at different speeds and a performance based training program would likely decrease training time while improving retention. A 1990 US Institute for Defense analysis of 47 studies comparing the use of interactive computer videodisc education with traditional classroom education found CBT more effective for knowledge and performance, with lower attrition rates and instructional costs and a 31% reduction in learning time [16, 17]. In a study of six controlled environments directly comparing traditional classroom instruction to equivalent interactive multimedia instruction at companies such as IBM and Xerox, CBI resulted in 56% more learning in 38-70% less time [18]. Other studies have shown CBT in business reduced training costs 40 to 85 percent compared with traditional training [19].

7.3 New Technologies and Resources for Skills Development

An example of a promising new technology is the virtual system that has been developed by EPRI for developing skills and measuring proficiency in ultrasonic testing. The virtual system, by providing electronically created substitutes for flaws, components, and transducers, offers potential advantages over the use of real systems for certain training and evaluation functions. These include more comprehensive measurement of operator performance, detailed and timely feedback to the operator of performance results, more options for the administration of training, and potential cost reductions in training and qualifying operators. A study of the effectiveness of flight simulator training for the military showed that "an hour of simulator time saves an hour of training time in an actual aircraft" [20].

Other available resources for enhanced NDE skills development include the computer-based, distance learning, interactive educational tools and materials that have been prepared by EPRI, the Iowa State University Center for Nondestructive Evaluation, and others. A principal advantage of these interactive tools and methods is that they enable each student to develop at his or her own pace. Without such tools, group instruction must necessarily proceed at the pace of the slowest student. After passing the performance-based test that clearly proves the understanding of the topic, the student can proceed to the next topic or level.

7.4 A Structured Approach to NDE Skills Development

The development and implementation of a structured approach to NDE skills development is proposed, consisting of the most effective combination of new technologies and tried-and-tested methods. Human performance studies have shown the importance of including certain strategies into training such as immediate feedback and the use of technique aid cards. Computer-based interactive materials can be employed for entry-level and also more advanced knowledge-based training. Virtual systems can be used in developing skills in the cognitive aspects of NDE tasks-those that require reasoning, interpretation, application of knowledge, and the development and testing of hypotheses. Practical aspects of NDE tasks can be learned and practiced in the actual job environment. The objective is to structure the skills-development experience in a manner that applies the available technology and resources to produce better qualified NDE personnel in less time. Recognizing that individuals learn at different rates, an important aspect of the structured approach will be to accommodate individual differences. Adequate proficiency testing will be required at key points in the process so that an individual can move from one level to the next when ready, but only when ready. This objective oriented approach to training has been adopted by the Accreditation Board of Engineering and Technology (ABET) [21]. Schools with engineering technology curriculums are given the choice to be evaluated for accreditation in accordance with these new criteria or the prescriptive conventional criteria. For example, instead of the survey team ensuring that the curriculum includes a specified number of hours of engineering classes, the validity of the program objectives is assessed and their effectiveness is measured by how many students pass the Engineer in Training test. Furthermore, the effectiveness of the mandatory continuous improvement system is also evaluated.

7.5 Implementation

To implement the change to a performance-based system will require revisions to applicable standards. For the nuclear power industry ASME Codes and Standards such as Section III, V, and XI will need to be revised along with ANSI/ASNT CP-189 and ASNT Central Certification programs. It is envisioned at this time that the training hours will be replaced by performance testing for each topic and completion of approved distance learning programs. The experience requirements could be replaced with virtual training experience and mentors documenting satisfactory completion of each applicable NDE skill for NDE personnel under their supervision. The skills would be contained in a detailed plan that represents all the skills necessary for an individual to perform their job. Recertification every 3 or 5 years by written examination should be replaced with annual practice detecting realistic flaws and non-relevant indications.

7.6 Conclusion

As indicated in this paper, the technical basis exists to justify the change to a performance-based training and qualification that will likely decrease the time necessary to satisfactorily certify NDE Personnel. IBM cut its annual training budget (over $1 billion) by $30 million by using CBT. It will also improve their capabilities and maintain their NDE skills. It would have the additional benefit of reducing a major reason for NDE personnel leaving the workforce and thereby increasing the recruitment of new personnel into the workforce.

8
PERFORMANCE DEMONSTRATION IN THE USA (PISC III)

This chapter includes excerpts from a paper presented at the 1995 Pressure Vessel Piping conference [22].

8.1 Objective

In the 1989 Addenda of the ASME Boiler and Pressure Vessel Code Section XI, Appendix VIII, the rules for qualifying UT systems (personnel, equipment and procedures) were published. Representatives from utilities organized the Performance Demonstration Initiative Steering Committee (PDISC) to implement Appendix VIII in response to the new requirements. The utilities collaborated on their resources to provide a unified approach that is economical and utility directed.

Nearly all of the samples contain flaws that have been fabricated and are being used for the demonstrations. Pipe samples removed from service containing IGSCC have also been included in the program. The results of these initial demonstrations are summarized in this section.

8.2 Procedures

Candidates have several options when participating in the program concerning which demonstrations they will attempt—IGSCC or non-IGSCC, austenitic or ferritic samples, single sided or double sided, automated or non-automated, and length or depth sizing. The RPV program has twenty three samples that represent boiling water reactors (BWR) and pressurized water reactors (PWR). They range in thickness from 6.25 inches (16 cm) to 11.25 inches (29 cm) and the nozzle specimens range in diameter from 2 inches (5 cm) to 42 inches (107 cm). In accordance with the Code requirements, the piping specimens include IGSCC, mechanical fatigue cracks, and thermal fatigue cracks along with non relevant flaws such as counterbores, weld roots, and weld crowns. The RPV specimens contain cracks, fabrication flaws, and a few notches.

8.3 Results

At the time this report was constructed, not all of the candidates have had an opportunity to participate in the performance demonstrations. Approximately one third of the 450 potential manual piping candidates have made at least one attempt. Two of the RPV vendors have participated and one automatic piping demonstration has occurred. Preliminary results from the demonstrations are:

8.4 Piping

A majority of the manual piping candidates intend to qualify for IGSCC austenitic or non-IGSCC austenitic detection, length sizing, and depth sizing. Therefore, a candidate needs to pass all six phases to be completely qualified for either IGSCC or non-IGSCC. The acceptance criterion for depth sizing is a .125 inches (3 mm) RMS error. PDI has elected to use a Root-Mean Square (RMS) error of .75 inches (1.9 cm) for acceptance criteria for length sizing. A Code revision to reflect this has been submitted. The entire length sizing demonstrations are being graded again using the RMS error criteria instead of the Code requirements of slope, correlation and critical miscall criteria. Figure 8-1 is a graph of the passing rates by candidate for detection only, detection with length sizing, and detection, length, and depth sizing.

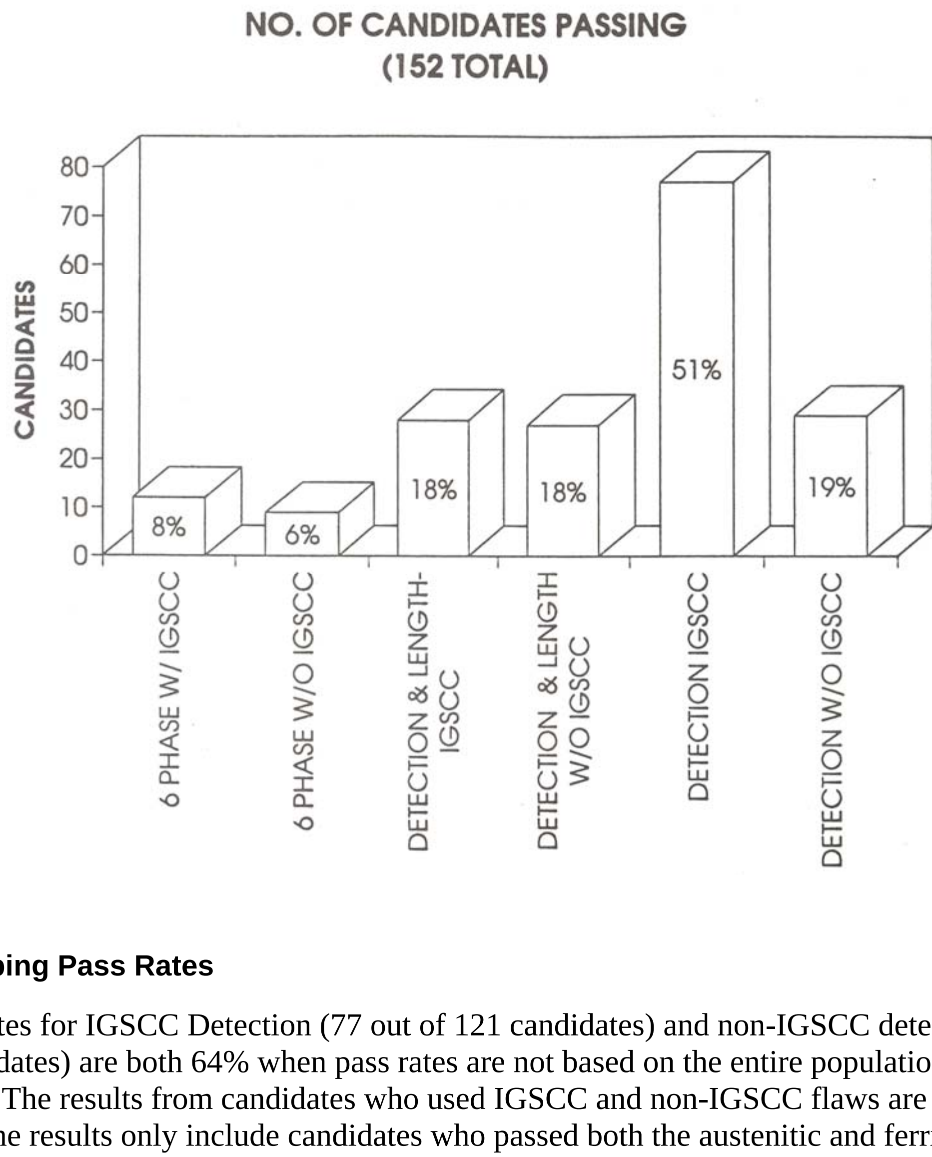


Figure 8-1
Manual Piping Pass Rates

The pass rates for IGSCC Detection (77 out of 121 candidates) and non-IGSCC detection (29 out of 45 candidates) are both 64% when pass rates are not based on the entire population of 152 candidates. The results from candidates who used IGSCC and non-IGSCC flaws are summed together. The results only include candidates who passed both the austenitic and ferritic portions of the demonstrations. Figure 8-2 is a graph of the number of successful attempts and total attempts made for detection and length sizing and then for all six phases (detection, length and depth sizing) of the demonstrations.

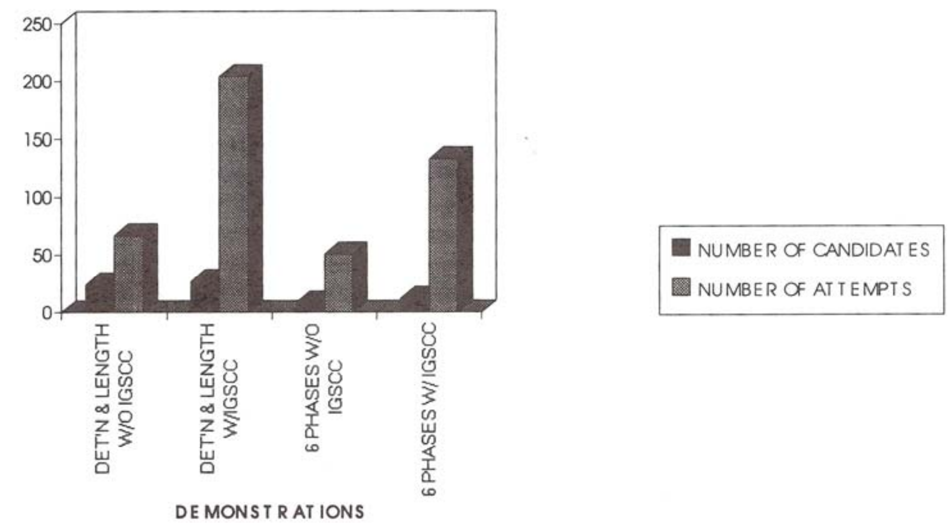


Figure 8-2
Passing Rates by Attempt

Figure 8-3 is a comparison of the passing rates for PDI's IGSCC program and the IGSCC detection program that has been conducted at the EPRI NDE Center since 1982 as part of the three party agreement between USNRC, EPRI, and the Boiling Water Reactor Owner's Group.

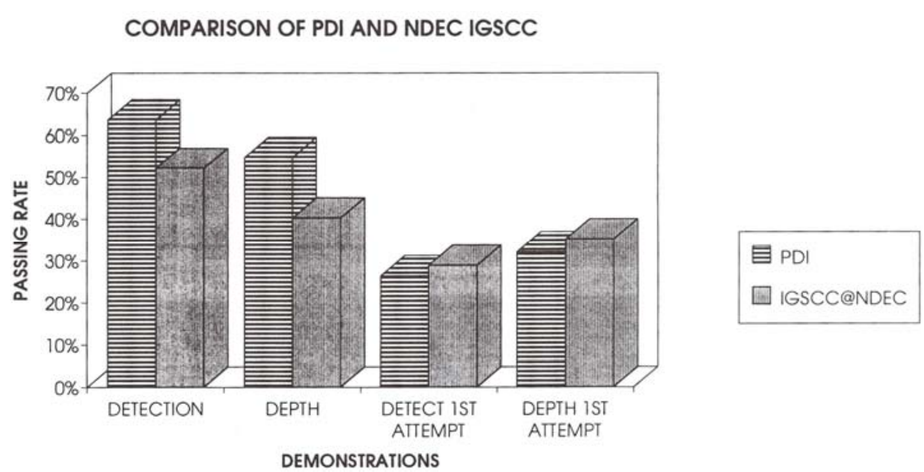


Figure 8-3
Comparisons of PDI and NDEC IGSCC

One team attempted the piping demonstration using an automated system and they did well. Three out of the four candidates passed detection and length sizing for a sample set containing IGSCC and ferritic specimens. Only one candidate passed the depth sizing demonstration using the automated system.

Note that the passing rate for the PDI IGSCC flaws (64%) is higher than the rate (52%) for the IGSCC detection course that began in 1982. The tests however are different. The PDI IGSCC program contains at least three IGSC cracks and three party agreement test contains ten IGSC cracks. The PDI program contains a larger variety of pipe sizes and ferritic specimens. Because of the differences in the test, the results are difficult to explain. The passing rate for the PDI's IGSCC depth sizing samples (55%) is also higher than the IGSCC planar flaw sizing course (40%).

8.5 RPV

At the time of this report, two vendors have attempted the PDI RPV demonstrations for Clad/Base Metal Interface Region and Reactor Vessel Welds Other than Clad/Base Metal Interface. They acquired the data and analyzed it during a six week period. Both of them successfully demonstrated their detection procedures and most of their sizing procedures. Figure 8-4 summarizes the results of the candidates who have taken the automated RPV demonstrations.

	Clad/Base Metal Interface			RPV Weld			Clad/Base Metal Interface			RPV Weld		
	Det'n	Length	TWE	Det'n	Length	TWE	Det'n	Length	TWE	Det'n	Length	TWE
Candidate 1	1	1	1	1	1	1	1	1				
Candidate 2	1	1	1	1	1	XX	na	na	na	na	na	na
Candidate 3	2	2	na	2	2	na	na	na	na	na	na	na
Candidate 4	1	1	X	1	X	X	na	na	na	na	na	na
Candidate 5	1	1	X	1	1	X	1	1	1	X	na	na
Candidate 6	1	1	1	1	1	X	na	na	na	na	na	na
Candidate 7	1	1	1	1	1	X	1	1	1	1	X	X
Candidate 8	1	1	1	1	1	X	na	na	na	na	na	na
Total Personnel	8/8 100%	8/8 100%	5/7 71%	8/8 100%	7/8 88&	1/7 14%	3/3 100%	3/3 100%	3/3 100%	2/3 67%	1/2 50%	0/2 0%
Total Procedure	1/1	1/1	1/1	1/1	1/1	1/2	1/1	1/1	1/1	1/1	1/2	0/2

Key:

1	Candidate had acceptable results on 1st attempt
2	Candidate had acceptable results on 2nd attempt (1st retest)
X	Candidate had unacceptable results on 1st attempt (eligible for retest)
XX	Candidate had unacceptable results on 2nd attempt (eligible for retest)
na	Supplement/application not attempted by the candidate
Det'n	Detection
TWE	Through Wall Extent (Depth)

Figure 8-4
Summary of Inside Surface Demonstrations

Initial results for demonstration performed from the inside surface are excellent except for depth sizing (TWE) of subsurface flaws in the RPV welds. It is not surprising to see that demonstrations performed only one side of the weld are not as effective as those performed when access is provided to both sides of the weld. The results for demonstrations performed from the outside surface Figure 8-5 are not as effective.

	Clad/Base Metal Interface			RPV Weld		
	Det'n	Length	TWE	Det'n	Length	TWE
Candidate 1	XX	1	1	1	1	1
Candidate 2	2	1	X	1	1	X
Candidate 3	2	1	X	X	1	1
Candidate 4	XX	1	X	1	1	X
Total Personnel	2/4 50%	4/4 100%	1/4 25%	3/4 75%	4/4 100%	2/4 50%
Total Procedure	1/1	1/1	1/1	1/1	1/1	1/1

Key:
- 1 Candidate had acceptable results on 1st attempt
- 2 Candidate had acceptable results on 2nd attempt (1st retest)
- X Candidate had unacceptable results on 1st attempt (eligible for retest)
- XX Candidate had unacceptable results on 2nd attempt (eligible for retest
- na Supplement/application not attempted by the candidate

Figure 8-5
Summary of Outside Surface RPV Demonstrations

8.6 Conclusions

PDI has developed a comprehensive program for implementing Appendix VIII in a timely, credible, and cost-effective way. All of the American nuclear units have joined the program, confirming the importance placed by utilities on NDE performance demonstration. The PDI program is organized and led by the utilities with technical and administrative support being provided by EPRI. The performance of the manual piping candidates has been a little better than the pass rates for the IGSCC detection and sizing programs conducted at the EPRI for the thirteen years before this report was written. This is to be expected since a majority of the candidates have been previously qualified at EPRI to detect and size IGSCC and the techniques and procedures have improved. For example, the generic procedures provide more guidelines on indication interpretation. The RPV vendors have been quite successful at detecting and length sizing flaws. There are no programs similar to PDI so a comparison of the pass rates cannot be made. The most difficult phase of the demonstrations appears to be the depth sizing of subsurface flaws.

9
PROGRAM FOR THE ASSESSMENT OF NDT IN INDUSTRY (PANI I)

9.1 Objective

A program of work has been undertaken to investigate the performance of in-service NDT used for standard key industrial plant components [23]. A management committee, involving senior managers from a broad range of UK industry, was established for the project to ensure that the results have the greatest possible relevance to current industrial concerns. A questionnaire was sent to industrial companies and resulted in manual ultrasonics being selected as the NDT method to be investigated by application on a number of test pieces in a round robin exercise. Test pieces which replicate key industrial components were produced containing artificial, service induced defects and these were mounted to simulated on-site access conditions. An ex-service boiler, which has been scrapped because it contained unacceptable defects, was also used as a test piece. The test pieces were inspected by teams drawn from UK inspection organizations, both inspection vendors and inspection departments within the industries concerned. The test pieces were sectioned to assist in the analysis of the effectiveness of the ultrasonic inspections. The results show a wide variation in the detection of the defects and in their sizing and positioning. The boiler was particularly challenging. The implications of the results and requirements for future developments are discussed.

9.2 Procedures

All the operators used standard ultrasonic procedures, which had been produced by AEA Technology and approved by the Management Committee, except for the following:

Company Z—Own procedure for boiler test piece P6
Company Y—Own procedure for all test pieces
Company X—Own procedure for test pieces 3 and 4
Company W—Own procedure for all test pieces

All the ultrasonic procedures were basically the same. They were all based on the now obsolete BS3923 Part 1 (9) and required sensitivity of 14 dB below a DAC (Distance Amplitude Curve) based on a 3 mm SDH (Side Drilled Hole). The probes were standard 0˚ compression and 45˚, 60˚ and 70˚ shear. The Company Y procedure stated that the application was for thicknesses greater than 10 mm (0.39 inch) and so it was not strictly applicable to the 7 mm (0.28 inch) nozzle in test piece P1.

The operators also used very similar equipment as might be expected. Seven used digital flaw detectors, while the remainder used analogue display sets.

All the teams were offered the opportunity to perform Magnetic Particle Inspection (MPI) on the test pieces prior to undertaking the ultrasonic inspection.

9.3 Results

The analysis of the results falls into two parts: the results obtained from the manufactured test pieces which contained known discrete defects; the results from the boiler which can only be compared to discrete destructive analysis. The manufactured test piece results allow for a general analysis of the performance of the operators in detecting the defects and this is presented below.

9.4 Manufactured Test Pieces

Overall, the manufactured defects showed good correlation between the intended defect size and angle and the size and angle observed in the sectioning. The interface between a ferritic weld and ferritic parent plate is transparent to standard ultrasonic testing and this allows manipulation of the implanted flaws. In austenitic materials, the weld/parent plate interface provides a change in acoustic impedance and causes problems with the implantation of flaws.

Removing the five defects detectable by MPI from the analysis for the ultrasonic inspections gives the ultrasonic operator detection rates listed in Table 9-1 and illustrated in Figure 9-1.

Table 9-1
Operator Performance in Detecting Defects with Ultrasonics in Manufactured Test Pieces

No of Operators	No of Defects Inspected with Ultrasonics by these Operators	No of Defects Detected by these Operator	Percentage of Inspected Defects Detected. (%)
1	20	6	30
1	20	7	35
1	20	8	40
1	20	9	45
4	20	11	55
1	20	12	60
1	18	11	61
3	20	13	65
3	20	14	70

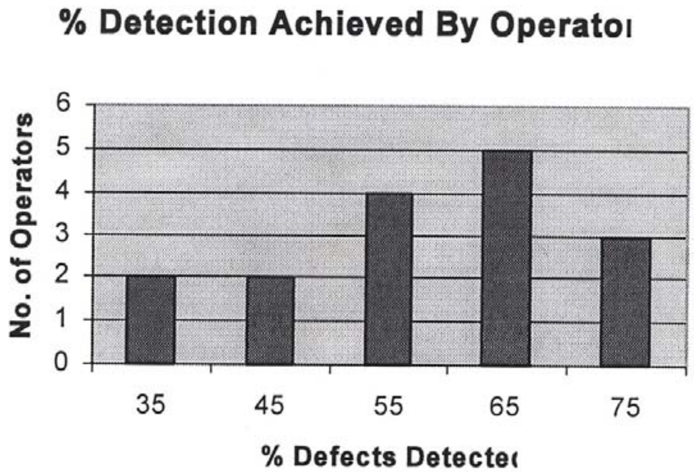


Figure 9-1
Operator Performance in Detecting Defects with Ultrasonics in Manufactured Test Pieces

The diagram gives data from 16 operators who were able to inspect most of the test pieces. One of the 16 did not inspect one test piece and his score was adjusted accordingly to give a percentage of the possible defects that he could have detected. Another of the 16 was only able to give a superficial examination of one of the test pieces, but his score has been left as a percentage of the total defect population. The highest detection rate of 70% was achieved by three operators: this equates to the detection of 14 of the possible total of 20 defects. Conversely, the low detection rates of 30% and 35% equate to the detection of 6 and 7 of the 20 defects respectively.

The detectability of individual defects is illustrated in Figure 9-2. False calls were few and appeared random.

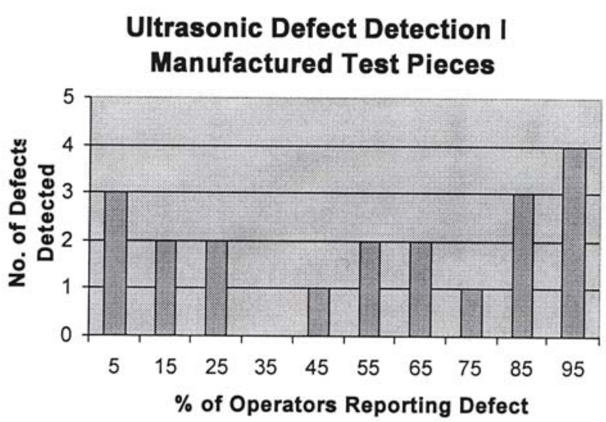


Figure 9-2
Detectability of Individual Defects in Manufactured Test Pieces

9.5 Discussion

Overall it may be concluded that defects were not detected for the following factors:

- The defect had a low response or a response which was hard to distinguish among other geometric echoes.
- The defect was only short in length and therefore missed if there was any lapse in concentration or gaps in the scanning raster.
- Human Factors. This is manifested by a multitude of factors including loss of concentration, poor scanning, poor coupling, and generally poor technique. Such effects arise from a variety of causes including poor motivation, poor training, adverse conditions for inspection, fatigue, and so on.

The detection rates achieved in the round robin, 30 to 70%, appear poor compared to the results of the PISC III Action 7 inspections which showed detection rates of 67 to 96%. However, the latter were obtained on a test piece 600 mm (23.62 in) by 500 mm (19.69 in) by 100 mm (3.94 in) with a double V weld prep. The defects were crack-like, vertical and parallel to the weld center line. They were 10 mm (0.39 in) to 15 mm (0.59 in) through wall and 20 mm (0.79 in) to 30 mm (1.18 in) long. They were therefore less difficult than the PANI specimens where some of the defect lengths were small compared to the probe width, or distracting geometric echoes were sometimes present. If this is taken into account, then the PANI results are consistent with the PISC III Action 7 results. These PANI results are in line with the NORDTEST conclusion that NDE may be a very unreliable and the Finnish observation that detection in simple butt welds is good but that detection in pipe joints, fillet joints, and stud welds is poor. The NIL-NDP detection rate of 50% is broadly consistent with PANI results.

9.6 Conclusions

Despite having good NDE qualifications, the operators found the inspections challenging. The detection performance of the operators was good on the relatively simple geometry test piece but deteriorated as the geometry increased in complexity. Overall the program showed that the application of a generic procedure, based on national standard, using operators with a generic qualification will not necessarily provide a good detection or sizing reliability particularly on complex geometries.

Overall the results were:

- Detection of individual defects was between 6% and 100%.
- Detection rates were not well correlated with size.
- Generally, far surface-breaking defects were most easily detected.
- Defects associated with geometrical features such as lands were hardest to detect.
- Two of the defects were below the reporting threshold

If the reporting criteria had been based on CEN 1714, four of the defects would not have been reportable.

10
PERFORMANCE DEMONSTRATION – 25 YEARS OF PROGRESS

This chapter includes excerpts from a paper presented at the 3rd International Conference on NDE in 2001 [24].

10.1 Abstract

Over the past 25 years, performance demonstration has progressed from an idea to a requirement that has been adopted by national and international codes and standards, and mandated by several national regulatory authorities. The progress has been slow as it progressed through several stages. These stages begin with recognition of the problem, usually as the result of some event that demonstrates the shortcomings of prescriptive examination requirements. We have now progressed through recognition, solution formulation, and initial demonstration programs. In this paper, the author discusses where we have been, what we have accomplished, problems encountered and questions for future consideration.

10.2 Introduction

The early 1970's saw a tremendous buildup of nuclear generation capacity. Pre-service and Iin-service inspections were new to the industry. The initial version of Section XI of the ASME Code was issued in January of 1970. The techniques used were adapted from manufacturing inspection. By today's standards, the equipment available for these examinations was crude with limited recording capabilities. The techniques relied on amplitude-threshold detection and amplitude-drop sizing techniques.

The discovery of a large defect during pre-service inspection at the E.I. Hatch plant raised a question as to the capability and adequacy of the examinations that were performed. Lacking confidence in the sizing capabilities of that time, the utility chose to repair the defect instead of calculating whether the defect was within acceptable safety guidelines. This even also stimulated the development of acceptance standards based on fracture mechanics that are now the basis for Section XI Acceptance Criteria.

Efforts to demonstrate the effectiveness of in-service inspection most often resulted in demonstrating its shortcomings. The Pressure Vessel Research Committee (PVRC) reactor pressure vessel (RPV) inspection trials and Plate Inspection Steering Committee (PISC I) are examples. These exercises further eroded the confidence in our NDE methods, in particular the ultrasonic method.

Other events that stimulated the need for improvement of inspection effectiveness included:

- IGSCC prompted NRC to shut down all operating BWR reactors for inspection in 1975.

- Thermal fatigue cracking found in BWR feed water nozzle; NRC required demonstration of effective inspection in REG Guide 06019.
- After the occurrence leaks in large diameter reactor coolant system (RCS) piping due to IGSCC, NRC required performance demonstration and qualification program in 1982.
- NRC issued Regulatory Guide 1.160 requiring improvement in RPV inspections in 1983.
- PISC II demonstrated improved detection; however, sizing capabilities were inadequate in 1984.
- Piping Round Robin demonstrated inadequate examination capabilities in 1982.

More recent events include:

- Dissimilar metal weld cracking in BWR units
- PWR RCS cracking at VC Summer and Ringhals
- Leaking Control Rod Drive (CRD) housing welds in PWR RPV closure heads

10.3 Solutions

Simply increasing inspection sensitivities was not a viable solution. The systems would be overloaded with information from harmless reflectors. It was clear that prescriptive requirements on examination techniques would not work for all applications. Individual solutions were required to solve particular problems. Performance demonstration was selected as the most appropriate solution. This scheme requires that particular procedure, equipment, and personnel combinations are capable of detecting and sizing flaws of concern.

10.4 Performance Demonstration

The ASME Code first adopted performance demonstration in the 1989 Addenda of the 1989 Edition. This Code is applicable to piping, bolting, and the RPV. The NRC has mandated accelerated implementation the Code requirements in a phased approach with the final components being completed by November of 2002. Since 1990, a large majority of the nations operating nuclear power stations have also adopted this approach or are considering some form of performance demonstration program. The European Network Inspection Qualification has published guidelines that are being adopted by individual european countries. The IAEA has also published a recommended methodology for Water-Water Energy Reactors (WWER). The large number of papers on this subject at this and the previous conferences are an indication of the movement toward performance-based approaches.

10.5 Progress

In the United States, PDI administers performance demonstrations for ultrasonic examination of the RPV, piping, and bolting. Each of the U.S. nuclear utilities and three foreign countries are members. The PDI was organized in 1991 to address the implementation of Appendix VIII performance demonstration requirements. Demonstrations were initiated in 1994 for piping, bolting, and RPV applications through Appendix VIII Supplements 2, 3, 4, 6, and 8. The first demonstration for weld overlay repaired welds (Supplement 11) was completed in September

2001. Initial demonstrations for dissimilar welds (Supplement 10) and RPV Nozzles (Supplements 5 and 7) are planned for completion by November 2002.

Since the time of PISC II, there have been tremendous advancements in computer capability and technology. These advancements have allowed for the development and improvement of automated examinations systems, particularly those used on the RPV. These improvements include both detection and sizing procedure enhancements. Figure 10-1 demonstrated the POD for flaws at the clad-to-base metal interface for examinations performed from the clad surface. Both passed and passed-plus-failed candidates are shown. The 95% confidence bounds are shown for the passed-plus-failed candidates. Every candidate including those that failed the examination detected every crack equal to or greater than 0.25 inch (6.35 mm). Figure 10-2 compares the sizing accuracy obtained by successful PDI sizing candidates to results from PISC II. The PISC results shown here are the results of the advanced techniques that ranged from a depth sizing standard deviation of 12 to 22 mm (0.47 to 0.87 inch). The PISC results included results from embedded as well as surface connected flaws. The ASME acceptance limit for sizing of embedded flaws is 0.25 inch (6.35 mm) which is well below the best PISC II result—0.47 inch (12 mm).

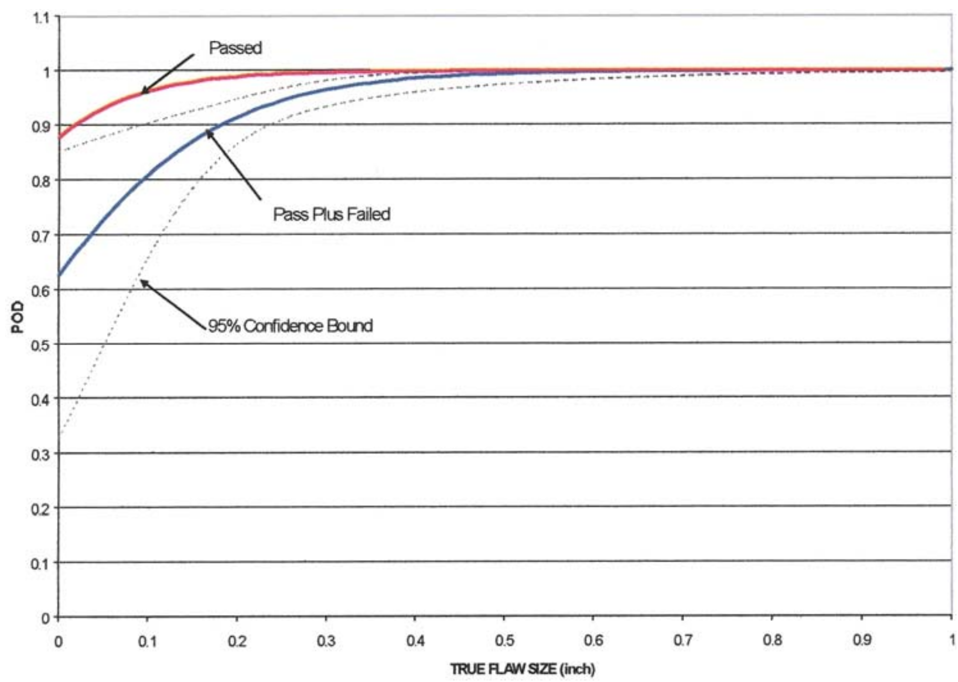


Figure 10-1
Probability of Detection for ID Examinations, for Flaws at the Clad-to-Base-Metal Interface, Based on PDI Demonstration Results

The results for passed and passed-plus-failed candidates are provided. The 95% confidence bounds are associated with the pass-plus-failed candidates.

Although not as dramatic, similar gains can be demonstrated for piping examinations. Figure 10-3 compares the POD for PDI piping candidates with those from three round robin exercises. These include the Piping Inspection Round Robin (PIRR) 1981-1982, the Mini Round Robin (MRR) 1986, and the PISC-AST 1989-1990. The data presented for PDI represents the same pipe sizes and flaw types as were included in the other three studies. The PDI data includes results from stainless steel piping with a wall thickness range of 11-25 mm (0.43 to 0.98 inch). Flaw types in the PDI data include thermal fatigue and field removed IGSCC. The PDI data include passed candidates and passed-plus-failed candidates over the years 1994 to 2001. Manual examinations make up the large majority of the PDI data included in Figure 10-3.

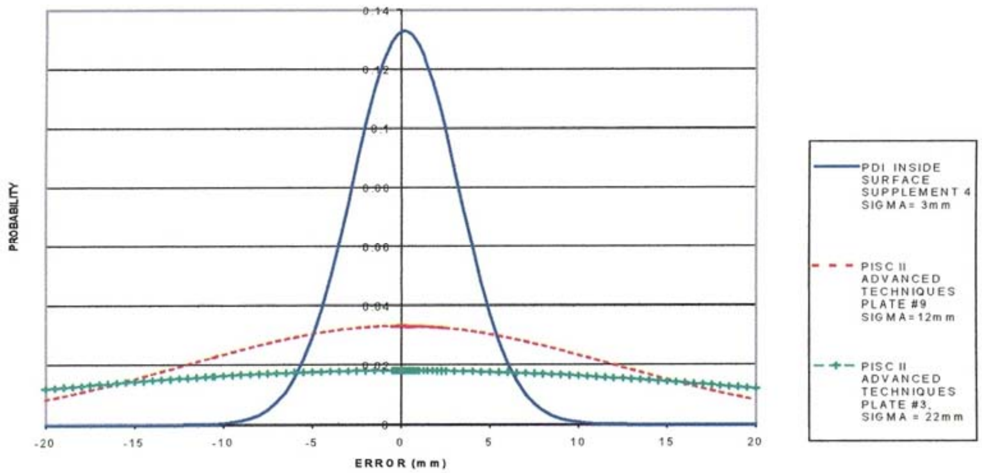


Figure 10-2
RPV Sizing Error Distributions for PDI and PISC II Advanced Techniques

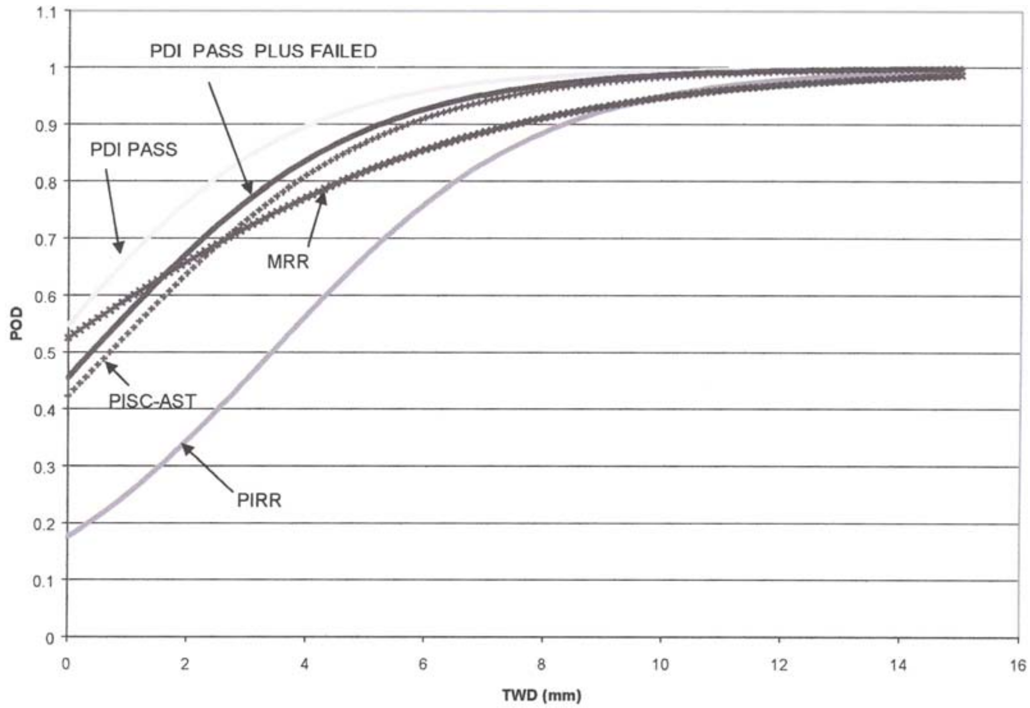


Figure 10-3
Comparison of PDI Candidate POD and Three Round Robin Exercises. Material is Austenitic Piping 11 to 25 mm (0.43 to 0.98 inches) in Thickness

There has been a steady improvement in the POD for austenitic piping examinations from 1981 to 2001. The overall improvement is substantial and may be attributed to better training and procedure improvements. Automated examinations have also benefited greatly from equipment improvements. These improvements in POD are accompanied by improved discrimination capabilities and fewer false calls.

10.6 Discussion

The industry has spent a large sum of money qualifying personnel and procedures. Many believe that this should solve all of our examination reliability problems. However, we are sometimes disappointed in the quality or effectiveness of examinations that are being performed. Performance demonstration/qualification is only one step in the total process. The process requires that the individual properly execute the procedures in the field, in the same manner they were qualified. The application must be within the scope of the qualification procedure, including access and geometric conditions. Qualifications alone are not a cure-all for the problem.

11
REACTOR PRESSURE VESSEL INSPECTION RELIABILITY BASED ON PERFORMANCE DEMONSTRATIONS

11.1 Background

Periodic in-service inspection of the RPV is performed using the ultrasonic examination method. These inspections are required on a 10-year interval. The presence of cracks at the inner surface of the RPV could be detrimental to the integrity of the RPV in the event of pressurized thermal shock or low-temperature over-pressurization events. Little or no credit has been given to the ability of ultrasonic inspection to identify flaws of concern to the integrity of the RPV. The lack of credibility demonstrated by early inspection reliability investigations contributed to this condition. In 1989, the ASME Code adopted a performance-based approach that included procedures, equipment, and personnel qualifications for ultrasonic examinations. U.S. utilities have implemented this requirement through the PDI 1007984 [25]. The EPRI NDE Program administers this program on behalf of the utility members. All U.S. utilities and seven foreign utilities are members of the PDI. Data collected from this program provide the basis for the examination effectiveness information provided in this report.

11.2 History

RPV ultrasonic inspection capability and reliability has been the subject of international round robin programs and laboratory studies for many years. One such program is PISC II, conducted in the early 1980s, which concluded that good flaw detection was possible but that defect depth sizing was inadequate. It soon became apparent that the prescriptive requirements of national standards such as the ASME Code could not provide adequate assurance that flaws of concern would be detected and sized as unacceptable. The direction chosen in the United States and much of the rest of the world was the adoption of a performance-based approach for qualification of ultrasonic examination systems, procedures, and personnel engaged in ultrasonic RPV and piping examinations. The ASME Code adopted the performance-based approach with the publication of Appendix VIII in 1989. Appendix VIII Supplement 4 addresses the RPV clad-to-base metal interface and the inner 10% of the vessel. Supplement 6 addresses the remaining 90% of the vessel thickness.

U.S. utilities organized and funded the PDI to address the implementation of Appendix VIII. All U.S. utilities and seven foreign utilities are members of PDI. EPRI was selected to administer the program at the EPRI NDE Center in Charlotte, NC. The program was initiated in 1991 with the first demonstrations taking place in 1994. Eight RPV plates weighing more than 56,000 lbs (25,401 kg) and containing more than 200 flaws were fabricated by PDI for Supplement 4 and 6 demonstrations. Twelve PWR and BWR nozzles were also provided for demonstrating nozzle inner radius and nozzle-to-shell examinations according to the requirements of Supplements 5 and 7. In 1999, the NRC mandated the implementation of Appendix VIII qualifications and specified a three-year phased implementation.

A measurement is defined as the interaction of one candidate with one flaw. Since the inception of RPV demonstrations, there have been 2,766 detection, 1,969 depth-sizing, and 3,840 length-sizing (total: 8,575) Supplement 4 and 6 measurements. The NDE Center has developed an inspection reliability database from this information, and this database provides information used in this report. A separate database for piping is also available. These databases provide insight into the capabilities and reliability of RPV examination as practiced in the United States.

11.3 Objectives

One objective of this report is to describe the methodology to measure examination performance based on performance demonstrations. A second objective is to use the available data to justify reduced inspection requirements based on the evidence of highly effective examinations. Estimates of RPV integrity or failure probabilities assume an initial flaws distribution. Effective examinations can reduce those distributions and minimize the need for repeated examination of low-risk areas of the vessel, for example, the outer 90% of the wall thickness. Examples of improvements in examination effectiveness are provided. These data will be valuable for use in probabilistic fracture mechanics (PFM) vessel integrity calculations.

11.4 Effectiveness of Examinations

Periodic in-service examinations are performed to ensure that the RPV is in satisfactory condition for continued safe operation. Flaws to be detected include initial manufacturing defects, manufacturing defects that have grown in service, and defects that have resulted from operation. The objective of performance demonstrations (ASME Section XI, Appendix VIII) is to ensure that the examination techniques are capable of detecting and sizing the target flaw types and sizes, that is, to ensure that they are effective. These performance demonstrations ensure this effectiveness through the use of a stringent blind procedure qualification. Personnel qualifications are performed to ensure that the individuals are capable of reliably executing the qualified examination procedures within acceptable norms of performance established in Appendix VIII. Both reliability and effectiveness are required to ensure that procedures performed by personnel and equipment are within acceptable bounds established by Appendix VIII.

All examinations have uncertainties. RPV examinations have both detection and sizing errors. Detection is essentially a binomial process-that is, the flaw is either detected or not detected. There can also be false calls. For example, the reported indication does not originate from a flaw. A candidate who has a large number of false calls reduces the confidence that the candidate could actually discriminate flaws from normal innocuous indications. Excessive false calls invalidate any PODs based on these data. Therefore, the number of false calls in all detection examinations is limited.

It is possible to estimate the probability that a flaw of a given depth will be detected. The estimate and its confidence bounds improve with the number of detection trials performed. There are additional uncertainties that can influence the POD. These include team-to-team and flaw-to-flaw variations. The PDI program addresses these variables by providing a conservative distribution of representative flaw types and collecting data from a large number of teams.

11.5 Passed-Plus-Failed Candidates

The large majority of candidates were required to detect 100% of the flaws in order to pass Supplement 4 and 6 acceptance criteria. The detection acceptance criterion allows one missed detection for test sets that have 12 or more flaws. Most candidates were tested on 7 to 10 flaws, which require 100% detection of the flaws in the test set. The Supplement 4 and 6 detection requirements are screening criteria as opposed to a method of measuring skill level over a continuous range of skill levels. The basis for the detection criterion was that no more than 5% of the unqualified candidates would be capable of passing the examination.

As an example, the POD for examination from the inside surface is provided in Figure 11-1. Curves are provided for both passed and passed-plus-failed candidates.

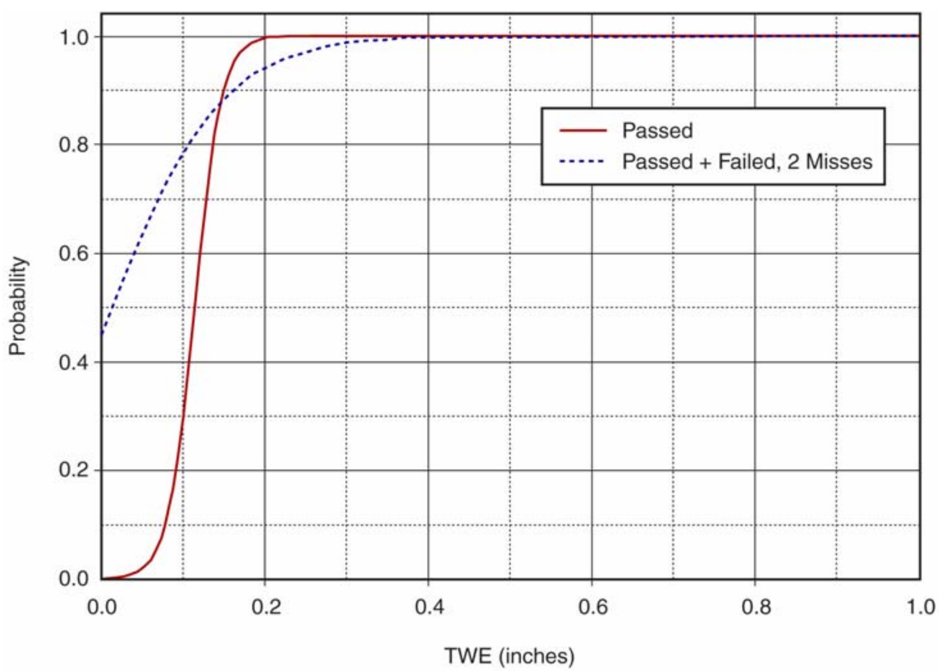


Figure 11-1
POD for Passed and Passed-Plus-Failed Candidates for Automated Examinations from the Inside Surface

Based on the results of the PDI demonstrations, Heasler calculated that the actual rate was closer to 1%. The calculations also showed that 5% or more of candidates that should be qualified would not be accepted.

Although 1% of the population might have passed by chance, it should not be a concern. However, to better understand the variability in performance of the process, a lower bound that describes the performance of a larger range of candidates is established. Potential choices for this lower bound include:

- the inclusion of passed and failed candidates with one missed detection
- the inclusion of passed and failed candidates with two or fewer missed detections
- the use of the lower confidence bound of the POD calculated for passed candidates

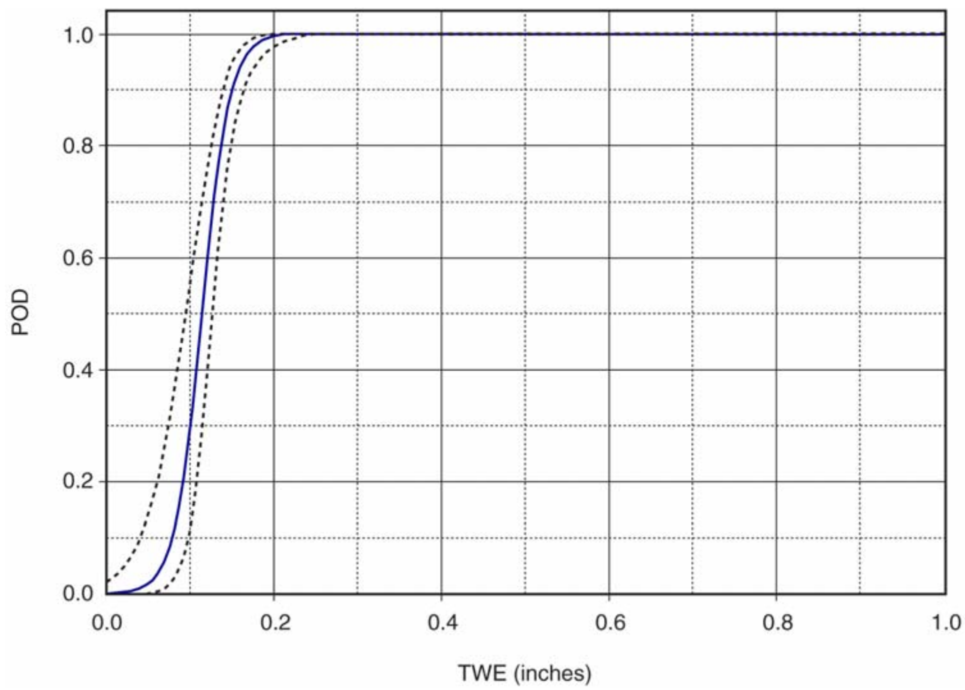


Figure 11-2
POD With 95% Confidence Bounds for Passed Candidates

The lower confidence bound for passed candidates is likely adequate to include passed candidates. However, with the large number of measurements, the lower confidence bound would be close to the estimated POD. The first two choices include a larger range of candidates. Candidates that failed but missed two or fewer flaws make up approximately 50% of the passed plus failed category for Supplement 4 measurements, that is, the inclusion of failed candidates doubles the population.

The passed plus failed category is included in many of the POD presentations. It should not be used as a representation of passed candidates. However, it does supply additional information regarding the broader population.

11.6 Probability of Detection

This section describes specific applications and comparisons of detection data. Examinations from the inside and outside surfaces of the RPV are addressed. Only automated procedures are used for examinations performed from the inside surface. The differences between automated and manual procedures are described for applications from the outside surface. Supplement 4 of Appendix VIII addresses flaws at the clad-to-base metal interface (the internal surface) of the RPV. Supplement 4 includes the inner 10% of the vessel thickness. This area is examined from the inside surface for PWR vessels and a limited number of BWR vessels and from the outside surface for the majority of BWR vessels. Supplement 6 of the Code addresses the remaining 85% of the inspection volume and includes imbedded and outside-surface connected flaws.

Flaws contained in the demonstration test specimens include crack like flaws that simulate manufacturing and postulated service-induced cracks. A sampling of volumetric defects simulating slag and slag combined with lack of fusion and cracking are also included. The range of flaw sizes is specified by the applicable Appendix VIII supplements. The exact size and distribution of the flaws in the PDI sample set are confidential. In most instances, the data presented here are extrapolated for a short distance beyond the actual flaw sizes to preserve the confidentiality of the test set.

Table 11-1 lists the number of detection measurements for eight combinations of supplements, accesses, and inspection methods. A total of 2,766 measurements are listed. This table lists the number of measurements by candidates who met the acceptance criteria of Appendix VIII passed, passed-plus-failed candidates, and the number of measurements excluded from consideration. The rational for these categories is provided in the report. The criteria for passed plus failed candidates are:

- No more than two defects were missed. This is equivalent to a minimum detection criterion of 70% as opposed to the Supplement 4 minimum acceptance criteria of 92%.
- No more than two false calls were made.

The column labeled “Excluded” is the total number of measurements less the number of passed plus failed measurements.

Table 11-1
Number of Detection Measurements

Supplement	Access	Auto/Manual	Passed	Passed + Failed 1 Miss	Passed + Failed 2 Misses	Excluded
4	ID	Auto	717	814	843	0
	OD	Auto	146		357	10
	OD	Manual	103		147	54
4	OD	All	249	287	504	64
Supplement 4 subtotal	**All**	**All**	**966**		**1347**	64
6	ID	Auto	658		783	101
6	OD	Auto	263		279	0
6	OD	Manual	73		122	63
6	OD	All	343		408	63
Supplement 6 subtotal	**All**	**All**	**1001**		**1191**	**164**
Total	**All**	**All**	**1967**		**2538**	**228**

11.7 Data Interpretation

The data provided in this section of the report are current as of January 2004. The information was collected earlier and illustrates the methodologies used to calculate inspection performance statistics. The data contained in this section as well as in Section 4 are extrapolated beyond the available data points on both the upper and lower ends in order to show the probable shape of the curve as well as to preserve the confidentiality of flaw sizes that are used in the PDI demonstrations. Interpretations of the data below a flaw size of 0.10 in. (2.5 mm) are not recommended.

12 COMPARISON OF PERFORMANCE TEST RESULTS WITH FIELD PERFORMANCE MEASURES

12.1 Background

A recent study presented at the 15th Steam Generator Workshop in 1996 [26] illustrated the comparison of performance demonstration measures with field performance measures under the following conditions.

- Eddy-current, bobbin-coil detection of ODSCC at tube-support locations
- 0.75 inch diameter steam generator tubes
- Seven categories of signal voltages for ODSCC indications

A sample of 818 steam generator tubes was examined by 12 Qualified Data Analysts (QDAs) under performance testing conditions, and separately, by 2 other QDAs under field conditions. The sample of tubes contained 1,363 reportable ODSCC indications at tube support locations. All analysts under both conditions worked independently of other analysts. A comparison of the average percentages detected under the two conditions is provided in Table 12-1.

The results showed a very high positive correlation between performance-test detections rates and field detection rates (a correlation coefficient of 0.999), even though performance measures were obtained from two separate groups of analysts rather than from the same analysts under each of the two conditions. Field detection rates were found to be just slightly less than performance-test detection rates, averaging 0.008 percent points less.

Thus, in this single, limited assessment, performance-shaping factors did not significantly degrade NDE capabilities under field conditions, and field performance could have been accurately predicted from measures obtained from performance demonstrations. On the other hand, the signal voltages of the ODSCC indications greatly influenced detection performance. Based on these results, one could not expect detection rates of 90% or greater unless signal voltages were equal to or greater than 1.00.

Table 12-1
Comparison Between Average Percentages of ODSCC Indications Detected at Steam Generator Tube-Support Locations from Performance Tests (Test) and Field Performance Measures (Field)

Signal Voltage	% Detected		Difference
	Test	Field	
<0.25	0.356	0.351	-0.005
0.25-0.49	0.566	0.535	-0.031
0.50-0.74	0.709	0.698	-0.011
0.75-0.99	0.820	0.820	0.000
1.00-1.49	0.928	0.936	+0.008
1.50-1.99	0.994	0.991	-0.003
≥2.00	0.993	0.976	-0.017

12.2 Conclusions and Recommendations

Potential contributions of aptitude testing to the selection of candidates for NDE jobs were illustrated by the development and validation of the Dynamic Inspection Aptitude Test (DIAT). This test was designed to predict success in the application of ultrasonic NDE techniques that require the combined aptitudes of general cognitive ability, abstract reasoning, and spatial visualization. Validation studies showed that the DIAT was highly reliable, correlated with other measures of these key aptitudes, and were predictive of the job performance of NDE operators. Moreover, they demonstrated that use of the test could result in considerable cost savings in the training and qualification of NDE personnel.

Effective NDE planning requires accurate measures of the capabilities of NDE systems, the combinations of techniques and operators to be employed in examining nuclear power plant components such as steam generator tubes, and pressure-vessel and piping components. To this end, performance tests have been developed and shown to be valid, reliable, practical methods for measuring the capability of NDE techniques and personnel [27, 28, 29].

However, measures of capabilities may not necessarily predict performance under field conditions, particularly if field conditions differ significantly from those under which performance testing is conducted. Numerous factors in the field environment can operate to degrade system performance. Consequently, research is required to compare the extent to which performance-shaping factors degrade NDE capabilities when they are applied under field conditions.

Once recent, but limited, study indicated that capabilities for the eddy-current bobbin-coil detection of ODSCC in steam generator tubes was hardly degraded at all under field conditions. However, more comprehensive studies of this type are required to provide a definitive basis for the prediction of field performance from measures obtained from performance testing. The goal of the recommended research would be to develop algorithms which can be applied to performance test results to predict NDE performance in the field.

13
REFERENCES

1. Ward D. Rummel. *Capabilities Data Book*, *Third Edition*. November 1997.
2. Ward D. Rummel; Steven J. Mullen; Brent K. Christner; and Robert E. Muthart. *Reliability of Nondestructive Inspection (NDI) of Aircraft Components*.   January 1984.
3. API Qualification of Ultrasonic Examiners Certification Program, American Petroleum Institute.
4. John Thompson. "PANI II Seminar Report." December 2003.
5. Jack Spanner, Jr. "White Paper Report 1: What is the Real Quality of ASNT Certified Personnel?" 2002 ASNT Fall Conference Panel. Unpublished.
6. Jack Spanner, Jr. "NDE Workforce Study", EPRI, Dec. 2000.
7. *Assessment of the IGSCC Training & Qualification Program.* EPRI, Palo Alto, CA: 1988. NP-5658.
8. F.L. Becker. "Performance Demonstration: 25 Years of Progress," 3rd International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components, November 2001.
9. P.G. Heasler and S.R. Doctor. "Piping Inspection Round Robin," NUREG, CR-5068, PNL-10475, April 1996.
10. P.G. Heasler, T.T. Taylor, S.R. Doctor, J.D. Deffenbaugh. "Ultrasonic Inspection Reliability for Intergranular Stress Corrosion Cracks: A Round Robin Study of the Effects of Personnel, Procedures, Equipment and Crack Characteristics," NUREG/CR-4908, PNL, July 1990.
11. P. Lemaitre. "Report on the Evaluation of the Inspection Results of the Wrought-to-Wrought PISC III Assemblies No. 31, 32, 33, 34, 35, and 36," PIDC III Report No 33, Programme for Inspection of Steel Components, Joint Research Centre EEC. September 1994.
12. J. Richardson. "API Personnel Proficiency Program-UT," Seminar on Advanced NDT Methods: A Technology Exchange, April 2001.
13. Jack Spanner, Jr. "White Paper Report 2: Proposed Performance Based Training and Qualification." 2003 ASNT Spring Conference Panel. Unpublished.
14. *Dynamic Inspection Aptitude Test User's Guide*. EPRI, Palo Alto, CA: 1996. TR-106304.
15. A. Edland and J. Enkvist. "Operator Performance in a Blind Test Piece Trial," *Materials Evaluation.* Vol. 59, No. 4, p. 531-536. (May 2000).
16. J.D. Fletcher. "Effectiveness and Cost of Interactive Videodisc Instruction in Defense Training and Education," IDA Paper P-2372. Alexandria, VA: Institute for Defense Analysis. 1990.
17. *Meeting NDE Workforce Requirements in the Nuclear Power Industry*. EPRI, Palo Alto, CA: 2001. IR-2001-002.

18. G.L. Adams. “Why Interactive?” *Multimedia & Videodisc Monitor, 10, 20*. March 1992.
19. B. Hall. "Return on Investment and Multimedia Training," *Multimedia Training and Internet Newsletter*. July/August 1995.
20. J. D. Fletcher, D. E. Hawley, and P. K. Piele. "Costs, Effects and Utility of Microcomputer Assisted Instruction in the Classroom," *American Educational Research Journal*, No. 27, p. 783-806. 1990.
21. http://www.abet.org/images/ASEE_Insitutional_Session.pdf
22. Jack Spanner, Jr., EPRI; Bruce Sheffel, Detroit Edison; Larry Becker, EPRI. “Performance Demonstration in the USA (PISC III) PVP, Vol. 317/NDE-Vol. 14,” *Effectiveness of Nondestructive Examination Systems and Performance Demonstration*, ASME. 1995.
23. H. Bainbridge, Magdalen House; G. Georgiou, The British Institute of Non-Destructive Testing;. B.A. McGrath. “Program for the Assessment of NDT in Industry (PANI),” AEA Technology. 1999.
24. F.L. Becker, "Performance Demonstration - 25 Years of Progress", 3rd International Conference on NDE In Relation to Structural Integrity for Nuclear and Pressurised Components, EPRI, Palo Alto, CA: 2001. 1007984.
25. F.L. Becker. *Reactor Pressure Vessel Inspection Reliability Based on Performance Demonstrations*. EPRI.
26. D.H. Harris. “Capabilities of Eddy Current Data Analysts to Detect and Characterize Defects in Steam Generator Tubes.” Paper presented at the 15th Steam Generator NDE Workshop, Long Beach, CA (1996).
27. Douglas H. Harris. *Personnel Testing and NDE System Performance*. Anacapa Sciences, Inc.
28. *PWR Steam Generator Examination Guidelines, Rev. 4: Vol. 1*. EPRI, Palo Alto, CA: 1996. TR-106589-V1.
29. D.H. Harris, S.D. Brown, G.L. Henry, and M.M. Behravesh. “Steam Generator NDE Performance Demonstration Assessment,” *Scientific and Engineering aspects of Nondestructive Evaluation. PVP-Vol. 257.* M.N. Srinivasan, ed. American Society of Mechanical Engineers, New York, NY. 1993.

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI)

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, California, and Charlotte, North Carolina, was established in 1973 as an independent, nonprofit center for public interest energy and environmental research. EPRI brings together members, participants, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of electricity generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents nearly 15% of EPRI's total research, development, and demonstration program.

Together…Shaping the Future of Electricity

1016969

ELECTRIC POWER RESEARCH INSTITUTE

3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com