Human Factors in Nondestructive Evaluation (NDE)

A Literature Review and Field Observations

3002010462

Human Factors in Nondestructive Evaluation (NDE)

A Literature Review and Field Observations

3002010462

Technical Update, November 2017

EPRI Project Manager

M. Dunlap

All or a portion of the requirements of the EPRI Nuclear Quality Assurance Program apply to this product.

YES NO

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

THE FOLLOWING ORGANIZATION, UNDER CONTRACT TO EPRI, PREPARED THIS REPORT:

Human Factors International, Inc.

This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.

NOTE

ACKNOWLEDGMENTS

The following organizations prepared this report:

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

Principal Investigators
C. Latiolais M. Dunlap R. Swain
M. Dennis T. Stafford

Under contract to EPRI:

Human Factors International, Inc.
1680 Highway 1, Suite 3600
P.O. Box 2020
Fairfield, IA 52556

Principal Investigators
A. Carreon C. Gaddy M. Polkosky

This report describes research sponsored by EPRI.

EPRI acknowledges the support of the following individuals and organizations:

QUEST NDE Services Inc. 19570 Meta Road Cornelius, NC 28031	Principal Investigator A. Conti
DCR Workforce 7295 NW Beacon Square Blvd., Suite 201 Boca Raton, FL 33487	Principal Investigators G. Gbemudu M. Leger
Istrouma Enterprises Inc. 493 Davis Farm Dr. Salisbury, NC 28147	Reviewer R. Smilie
Bundesanstalt für Materialforschung und –prüfung (BAM) (Federal Institute for Materials Research and Testing) Unter den Eichen 87 12205 Berlin, Germany	Reviewer M. Bertovic

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

Human Factors in Nondestructive Evaluation (NDE): A Literature Review and Field Observations. EPRI, Palo Alto, CA: 2017. 3002010462.

ABSTRACT

The job of inspecting nuclear plants represents one of the most extreme working environments possible from a human factors perspective. Examiners perform their jobs under conditions of high physical and psychological stress from a myraid of sources, including heat, radiation exposure, time pressure, noise, and the possibility of a catastrophic radioactive leak if they fail in their task. These variables induce substantial sensory, physical, and cognitive load on examiners. Therefore, it is important to consider the human factors inherent in the nondestructive evaluation (NDE) process and how they might be better addressed to improve examiner performance and NDE reliability.

To date, research has acknowledged the stressful nature of the NDE examiner's role. However, there are gaps between the research and actual practice. Based on this review of the literature, field observations, and review of operating experiences, human factors have much to offer in relation to NDE reliability, including psychological knowledge (such as stress, performance, and training) and applied research methodology. Human factors practices—including on-site observation, usability evaluation, user-centric design of equipment, and state-of-the-art training design—play a critical role in increasing the safety, satisfaction, and performance of NDE. NDE simulation appears to be a promising means of better preparing examiners for the incredible challenges of their difficult but critical role in maintaining public safety.

Keywords

Cognitive load
Human factors
Nondestructive evaluation (NDE)
Stress

Deliverable Number: 3002010462

Product Type: Technical Update

Product Title: Human Factors in Nondestructive Evaluation (NDE): A Literature Review and Field Observations

---

PRIMARY AUDIENCE: Researchers and NDE managers

SECONDARY AUDIENCE: NDE examiners

KEY RESEARCH QUESTION

What is the extent of human factors research within the nuclear NDE industry as it pertains to manual ultrasonic testing (UT)? Additionally, manual UT human performance information will be gathered from a review of operational experience within the nuclear industry and on-site visits to nuclear power plants. This information will be used to help guide future NDE human performance research.

RESEARCH OVERVIEW

This report reviews the available research and operating experiences of human performance in nuclear NDE, with an emphasis on manual angle beam UT weld inspections. It also describes on-site observations of NDE examiners. These activities showed that there is a limited number of independently verified studies for identifying the human performance variables that have a positive or negative effect on NDE reliability. Furthermore, there are multiple observable differences between training and practice settings; the implications of these settings provide a significant opportunity to apply existing and new human factors psychological knowledge and research methodologies to better prepare examiners for efficient, effective, and satisfying nuclear NDE.

KEY FINDINGS

- Nuclear NDE human performance research literature can be improved with additional up-to-date human factors research.
- In the past decade, very few ultrasonic NDE errors in which a qualified procedure performed by qualified personnel failed to correctly implement the procedure have been identified.
- First-hand examiner experiences and perceptions are rarely considered or measured.
- Future human factors research plans are covered.

WHY THIS MATTERS

Based on this review of the literature, field observations, and a review of operating experiences, human factors research and practice have much to offer for understanding and addressing NDE reliability. Specifically, we recommend that common human factors practices—such as on-site observation, usability evaluation, user-centric design of equipment, and state-of-the-art training—play a role in increasing the safety, satisfaction, and performance of NDE. Ongoing development of NDE UT simulators, which addresses several of these needs, appears to be a promising approach to better preparing examiners for the challenges of this difficult but critical role.

HOW TO APPLY RESULTS

The results can be used to better understand the differences and implications of laboratory training and certification of NDE examiners and the typical operating environment of a nuclear plant. These results will help prioritize and guide interventions to lessen the negative impacts of extreme work settings on NDE examiner performance in the field or during training.

LEARNING AND ENGAGEMENT OPPORTUNITIES

- For this research to be widely adopted, collaboration between and review by EPRI members, industry leaders, and regulatory bodies must occur.

EPRI CONTACTS: Myles Dunlap, Technical Leader, mdunlap@epri.com

PROGRAM: Nondestructive Evaluation Program, 41.04.01

IMPLEMENTATION CATEGORY: Reference

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

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Nondestructive evaluation (NDE) in the nuclear power industry represents one of the most extreme and taxing work roles and work environments for skilled personnel. Examiners perform their jobs under conditions of high physical and psychological stress from myriad sources, including heat, radiation exposure, time pressure, noise, and the possibility of a catastrophic radioactive leak if they fail in their task. These variables induce substantial sensory, physical, and cognitive load on examiners.

Human factors in NDE, specifically for ultrasonic testing (UT), have been studied by a variety of organizations and industry leaders. However, the scope of such studies has been limited and lacks a significant amount of quantitative evidence. Importantly, many previous studies have been conducted in the controlled environment of a laboratory. This raises the question of how well these studies reflect an individual's performance during the more challenging physical and social environment of field NDE.

Dozens of performance-shaping factors (PSFs) have been postulated to influence NDE reliability in the field (see Appendix A). Many of these, although of theoretical interest, might not be actionable in practical ways when an examiner is conducting field NDE. In the context of this report, NDE reliability will be considered as the degree to which the results of an NDE observation (that is, measurement, calculation, exam conclusion, and so on) can be depended on to be accurate.

For example, it is well known that individuals have different cognitive styles that could be reasonably expected to impact their NDE performance. Therefore, describing this variability is an important part of the theoretical literature on human factors in NDE. On the other hand, controlling or changing cognitive styles across all examiners is an impossible undertaking.

Researchers and practitioners simply cannot eliminate human variability, which means it will continue to exert a non-zero impact on NDE reliability in the field. In other words, as desirable as it might be to control a specific examiner's stress response, some of the distractions, decision-making skills, and other personal factors, the inherent variability in human cognition and behavior will decrease NDE reliability across all examiners.

Therefore, human factors research must characterize the most critical, actionable factors that influence examiner performance in a field NDE environment. First, this report focuses on a due diligence review of existing research summarizing the available evidence on NDE reliability and describes the two major NDE environments (training and field). The gaps in this research are identified, and a list of potential solutions is presented. From this foundation, a review of operating experiences and observations of in-field NDE is described, followed by a summary of major conceptual models that have guided research attention through the last decades. Finally, the report culminates in a proposal for improving the performance and job satisfaction of NDE examiners, which aims to improve NDE reliability. Moving toward a human factors approach to NDE reliability will require the involvement of and collaboration with industry leaders and regulatory bodies.

2
LITERATURE REVIEW

This literature review focused primarily on the use of manual angle beam UT weld inspection in the nuclear industry. Manual UT can be conventional (single beam angle) or phased array (multiple beam angles that form an image). For the most part, this report does not evaluate encoded or automated ultrasonic methods that record data.

NDE in the nuclear industry is also characterized by two very different environments—the laboratory environment and operating (field) environment. The laboratory environment, the primary setting of training, is a typical office-type environment with flawed samples used for examiner practice. The operating (field) environment, the primary setting of NDE in a nuclear power plant, is characterized by extremes of temperature and stress, radiation exposure, use of the required UT instrument (may be different than that used for evaluation and practice) and protective clothing, and physical challenges involved in accessing and examining welds.

In contrast to most other industries, NDE in the nuclear operating environment involves exposure to multiple interacting variables that can be expected to negatively impact an examiner's performance. These variables are also unique to each weld examined, even in a single nuclear power plant; although one might be easily accessible, another might be in a difficult-to-reach location that challenges even the most experienced examiner in positioning one's body and the evaluation equipment.

At the same time, the examiner is pressured by time constraints, heat, radiation, noise, and other factors. Also, different welds (material types and configurations) require different procedures, equipment, setups, and mindsets for examination.

Finally, NDE examiners often work for contracting organizations and travel to and from several nuclear power plants during the outage season. So, for any given examiner, the sequence of nuclear facilities; their site-specific differences; and the number, difficulty, and pace of exams—combined with the stressors of traveling, long work hours, and the physical and psychological setting of each plant—contribute to a stressful work environment.

The Yerkes-Dodson Law describes the relationship between arousal (stress) and performance as an inverted-U function, which indicates that as arousal increases, performance also increases to an optimum level then begins to deteriorate. The stressful environment of nuclear NDE will likely result in performance deterioration in potentially predictable ways. Under conditions of overarousal, as can be expected in nuclear NDE, individuals could experience the following, according to Wickens, Gordon, and Liu [1]:

- **Perceptual or attentional narrowing (tunneling):** a restricted range of attention while possibly ignoring other potentially relevant information
- **Cognitive tunneling:** focusing attention on one hypothesis and ignoring a potentially better diagnosis by considering other options
- **Memory biases:** relying on the most overlearned skills and thoughts
- **Strategic shifts:** taking immediate or fast action, while potentially sacrificing accuracy

The Yerkes-Dodson Law applies to NDE examiners, who experience both psychological and environmental stressors as well as significant variability within and between each NDE they perform, all of which increase the stress that they experience in a single day on the job. The literature review covered in the following sections will focus only on articles relevant to nuclear power plant applications of NDE, with a focus on manual UT, due to the unique challenges it presents for optimal human performance.

Performance-Shaping Factors

Variables thought to influence NDE reliability are known as *PSFs*. Dozens of PSFs have been postulated to influence NDE reliability. Swain and Guttman [2] provided one of the earliest catalogues of PSFs. The external PSFs that they identified can be found in Table 3-2 of [2] and are as follows:

- **Situational characteristics** or PSFs general to one or more jobs in a work situation, including architectural features, quality of environment (temperature, humidity, air quality, radiation, lighting, noise, vibration, cleanliness), work hours, shift rotation, availability/adequacy of special equipment, manning parameters, organizational structure, actions by supervisors and others, and rewards and recognition
- **Job and task instructions**, including procedures required, written or oral communications, cautions and warnings, work methods, and plant policies
- **Task and equipment characteristics**, including perceptual requirements, motor requirements, control-display relationships, anticipatory requirements, interpretation, decision-making, complexity, narrowness of task, frequency and repetitiveness, task criticality, long- and short-term memory, calculation requirements, feedback, dynamic versus step-by-step activities, team structure and communication, man-machine interface factors

Stressor PSFs include the following two types, which can be found in Table 3-2 of [2]:

- **Psychological stressors** such as suddenness of onset, duration of stress, task speed, task load, high jeopardy risk, threats, long uneventful vigilance periods, conflicts of motives about job performance, absent or negative reinforcement, sensory deprivation, distractions, and inconsistent cueing
- **Physiological stressors** include duration of stress, fatigue, pain or discomfort, hunger or thirst, temperature extremes, radiation, atmospheric pressure extremes, oxygen deprivation, vibration, movement constriction, lack of physical exercise, and disruption of circadian rhythm

Finally, internal PSFs include **organismic factors** (characteristics of individuals) that result from internal and external influences, including previous training/experience, state of current practice or skill, personality and intelligence variables, motivation, attitude, emotional state, stress, knowledge of required performance standards, gender differences, physical condition, attitudes based on influence of family and others, and group identifications [2]. The behavioral science researchers, studying ultrasonics within NDE, have shown a broad acknowledgment of the numerous variables impacting personnel in nuclear power plants, which is consistent with the broader human factors literature and shows the many performance stressors considered by research.

Literature Review Methodology

A statistical meta-analytic approach to the literature review was originally planned. Meta-analysis is a customary approach for quantitatively summarizing the data of a large body of research. As such, meta-analysis requires a group of studies that focus on common independent and dependent variables and use the same statistics. Next, these statistics are quantitatively summarized to determine overall trends in the literature.

An independent variable is a controlled or observed variable that influences another variable, called the *dependent variable*. For example, to meta-analyze the relationship between cognitive style (independent variable) on NDE performance (dependent variable), several studies would be required that define and quantitatively measure both of these variables and provide inferential statistics. In turn, these statistics are then further analyzed to calculate effect sizes for the entire body of research and arrive at an interpretation.

However, NDE research has not addressed similar variables repeatedly, nor have most studies been quantitative, which prohibits a formal meta-analysis of this literature. Further, the relatively small number of studies on the same independent variables, combined with the lack of repeated empirical investigation and quantitative support, is typically viewed as methodologically weak in the behavioral sciences. The current state of the literature required a revision to the planned methodology.

Thus, the current methodology used a content review and summary of the existing research. In keeping with a meta-analytic approach, this review includes studies that address the effect of various independent variables, or PSFs, on NDE reliability. Studies that were not primarily concerned with a relevance to manual UT in the nuclear power plant industry were not included, an approach that substantially limited the sample of relevant literature.

Review Summary

The selected high-quality and nonrepetitive (that is, generally not already published elsewhere) literature consisted of 24 research articles as described in Appendix A.

One research type was assigned for each article based on the best category:

- **Empirical:** based on, concerned with, or verifiable by observation or experience rather than theory or pure logic
- **Survey:** a general view, examination, or description of someone or something
- **Theoretical:** concerned with or involving the theory of a subject or area of study rather than its practical application

The publication type for each article was described as one of the following:

- Book chapter
- Conference paper/proceedings
- Dissertation
- Journal
- Technical report

Table 2-1 shows the research types, and Table 2-2 shows the publication types.

Table 2-1
Overview of research types analyzed in Appendix A

Research Type	Count
Empirical	20
Survey	3
Theoretical	1

Table 2-2
Overview of publication types analyzed in Appendix A

Publication Type	Count
Book chapter	1
Conference paper/proceedings	7
Dissertation	1
Journal	2
Technical report	13

The empirical basis for most of these articles—that is, verifiable by observation or experience rather than theory or pure logic—is helpful for practical applications. In this literature, the PSFs that received the greatest attention were procedures and training. Recommendations were made to improve procedure utility and usability and training fidelity. Other PSFs considered of importance in at least two or more of the studies were “environmental” stresses (physical, such as heat and time pressure), examiner experience, human-machine interface, supervision, team performance, and work processes, such as flaw detection.

Most of the articles were published as conference papers or technical reports, and as a result, they might be harder to find than a more widely published journal article or book. Also, many of the articles are older, and some are based on work done in other countries/languages, which can make finding them even more challenging.

Due to the technical focus of this body of work and its multidisciplinary nature, one can expect to find more articles in technical reports. In fact, technical reports from authoritative sources, such as research institutes and regulatory bodies, might be the best way in which to communicate with the intended audience. Many of the NDE journals are focused on technical aspects, and many of the psychology and related social science journals are focused on more generalizable results.

3
CONCEPTUAL MODELS

A conceptual model describes the theoretical relationships among variables. Not only do conceptual models assist with understanding, but they provide a mechanism for empirically testing human factors theory.

In the case of human NDE performance, conceptual models provide understanding of the many PSFs that impact an examiner's performance. In psychology, conceptual models have also evolved to provide a basis for empirical, quantitative examination of theoretical relationships using complex regression and advanced multivariate statistical methods.

One of the first influential conceptual models of NDE performance was advanced by Spanner [3]. This model specifies two major organizational variables (management and training) and a variety of perceptual and cognitive variables thought to causally influence an examiner's action and, finally, required data recording in the NDE procedure. In addition, information and procedures have bi-directional influence on examiner action. The model acknowledges the role of the machine as well as the environment, although directional variable relationships are not specified. This model is a reasonable starting point (especially given the state of psychological science at the time of publication), but it is dated in its conceptualization of causal relationships among variables and does not specify an outcome that is specific to human accuracy or group reliability of NDE performance.

The Harris model [4] (see Figure 3-1) is a somewhat simpler approach to describing variables involved in NDE performance and uses one of the major psychological movements at the time—reinforcement theory. At this point in psychology, research was dominated by theories that addressed primarily how individuals learn through reinforcement and feedback, often excluding internal, unobservable factors. In this model, Harris postulates that both information and PSFs influence the examiner, which in turn causes behavior and results in a feedback loop that promotes adaptive behavior. This conceptualization allows for the broad variety of PSFs and emphasizes the learning that occurs as an examiner gains experience. However, the model does not articulate the relationship between individual-level behaviors and knowledge and group-level variables, such as NDE reliability. It also does not differentiate among the many PSFs that might influence either examiner or NDE outcomes.

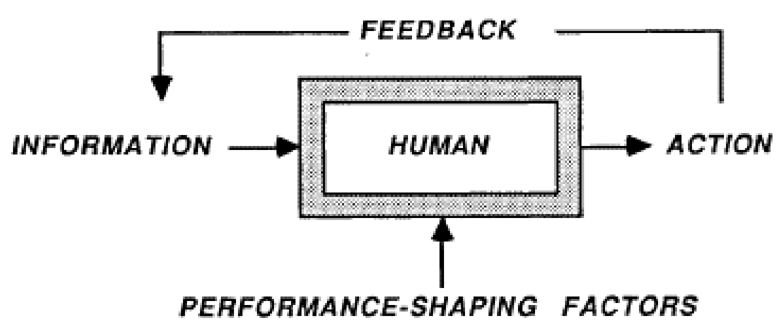


Figure 3-1
Harris model as shown in [4]

The Bertovic, Gaal, Mueller (also referenced in literature as Müller), and Fahlbruch model of NDE performance takes a broader approach to complex interrelationships among variables [5]. This model begins with a variety of perceptual, cognitive, and even social variables that are thought to influence mental workload and, in turn, performance. Technology and working conditions mediate the relationship between individual factors and mental workload. The organizational context is acknowledged as influencing organizational working conditions, indirectly influencing examiner performance. Finally, the relationship between an examiner's mental workload and performance is also mediated by stress reaction and satisfaction.

The commonality among these models is that they attempt to incorporate the multitude of variables that might influence examiner performance during NDE, while incorporating state-of-the-art psychology when they were published. As conceptual modeling matured, the models began to increasingly specify reciprocal, interacting, and mediating effects among variables, giving them greater descriptive fidelity for in-field NDE. However, despite the clarity and valid complexity of these models, even the most modern frameworks do not differentiate between actionable variables (those applied that researchers can improve, such as training, equipment usability, and some aspects of the working environment) and those factors that decrease human performance but are a required part of NDE in the nuclear industry (for example, physical working conditions, individual cognitive differences, and supervision and management differences).

Because a conceptual model serves as a guide to applied research, it is important to identify variables that will have focus in any research program. As shown by previous models, the interrelationships and number of independent variables impacting examiner performance in NDE are complex and difficult to represent visually because there are many bi-directional and interacting effects. For this reason, a linear causal model appears inadequate to describe the variables involved.

Figure 3-2 shows a proposed socio-ecological systems model of NDE drafted by the authors of this report, where examiner performance is the center of focus and categories of variables that have an impact on an individual examiner are shown in concentric circles. This approach to modeling is based on other conceptual models in psychology (for example, the Bronfenbrenner model) that have successfully articulated complex, interacting categories of variables. In this model, variables hypothesized to have a larger effect size are shown closer to the center (proximal variables), while those that are more distally related to examiner performance appear in outer rings. Thus, the independent variables that impact NDE examiner performance are as follows:

- **Extreme work environment**. This includes the physical environment, which might involve heat, radiation, humidity, physical positioning of examiner and/or equipment relative to a weld, time pressure, and the number and pace of exams completed in a given time span by an individual. The model acknowledges the atypical and extreme work environment of a nuclear power plant as the most critical category of independent variables, with powerful effects that have a strong primary impact on an examiner and likely interact with, and even exceed, other variables with smaller effect sizes.
- **Equipment design**. This level includes the display, controls, and method of operation of equipment required for NDE as well as any other personal protective or safety equipment required (that is, protective clothing, gloves, headgear, eyewear) that could interfere with an

examiner's vision or restrict his or her freedom of movement. The physical design of a nuclear power plant, including the accessibility of a given weld, is also included in this level because plant design will have a significant effect on an examiner's ability to perform NDE easily and accurately.

- **Training and procedures**. This level includes the specific procedure to be carried out, any documentation of that procedure, and variability in procedure for different welds within one plant. Also included here are differences in administrative procedures across multiple plants because many examiners travel among plants during an outage season and have an ongoing process of adaptation to their current environment. In addition, an examiner's training, ongoing skill maintenance, and levels of experience and certification are variables likely to impact his or her performance.
- **Organizational context**. Included here are variables related to work expectations, management, support of examiners, and organizational culture. Although these variables have received significant attention in recent years, a socio-ecological model views these variables as having a relatively small effect compared to more proximal variable categories.
- **Regulatory environment**. The overall set of factors thought to impact examiner performance includes the general culture of regulation and legal requirements governing examinations, particularly the safety culture and accountability to regulatory bodies. Over time, and depending on how recently missed flaws have been discovered or a more serious accident has occurred, the regulatory environment may be expected to change, which will naturally impact the organizational context of nuclear plants.

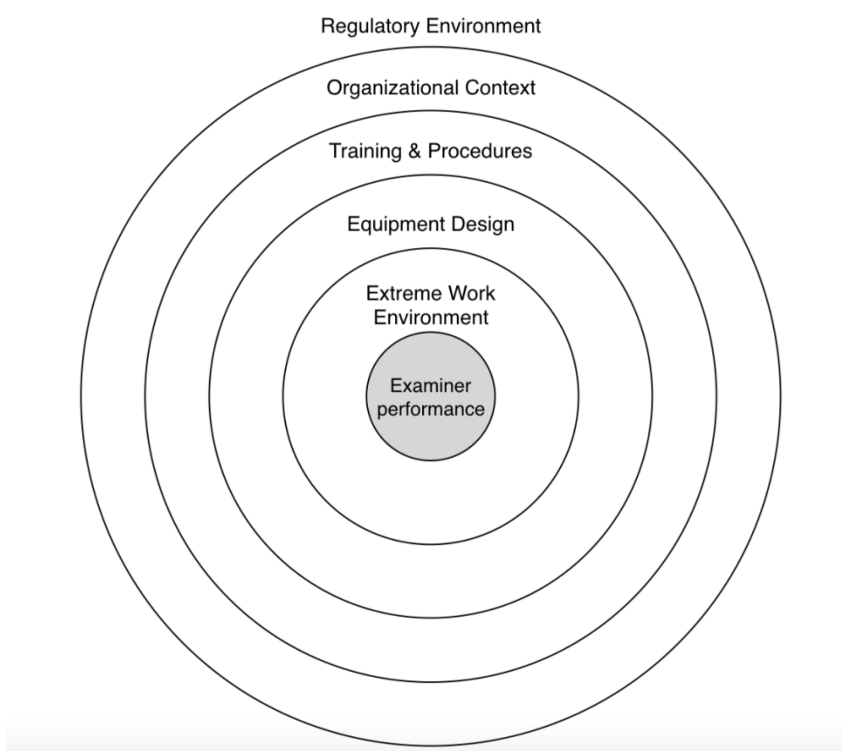


Figure 3-2
Socio-ecological systems model of an NDE examiner's performance

It should be noted that all these levels will interact for a single examiner: as one travels among several plants during outage season, in a strict regulatory environment, one may find strong but variable safety cultures with different procedures in each plant where the examiner performs NDE. In addition, each weld inspected provides a different, varying immediate environment, all of which might be considered extreme in the physical and mental challenges that they induce during an examination. Thus, for a single examiner, all five levels of variables might interact differently for a given examination, in a single plant, at a given moment in time. These effects will then change for the next examination, creating another source of cognitive and physical stress, as the examiner progresses through a single outage season.

Figure 3-3 proposes a socio-ecological model illustrating the impact of group performance on NDE reliability. As shown in Figure 3-3, each set of concentric circles represents an individual responsible for NDE reliability. The NDE reliability of a nuclear power plant is dependent on numerous individuals, each with different job roles and each role having its own unique set of influencing variables. Therefore, every individual’s contribution is different to NDE reliability. The proposed model in Figure 3-3 suggests how each individual contributor has his or her own set of influencing variables that affect NDE reliability. Figure 3-3 proposes that individual and group variables be cast similarly because NDE reliability is a function of how all personnel perform across all inspections. NDE reliability is not an individual level variable but a group variable. It is also significantly complex and all but impossible to measure, given the number of individual inspections that occur in even a single outage season. Further, because the true number of missed flaws is unknown, even with precise measurement of all inspections conducted in a specific period, NDE reliability is, at best, an approximation of the actual performance of examiners. It is important to recognize the limitation of NDE reliability as a sole outcome measure of examinations and develop additional outcome measures that can provide a more complete view of examiner confidence and performance in NDE.

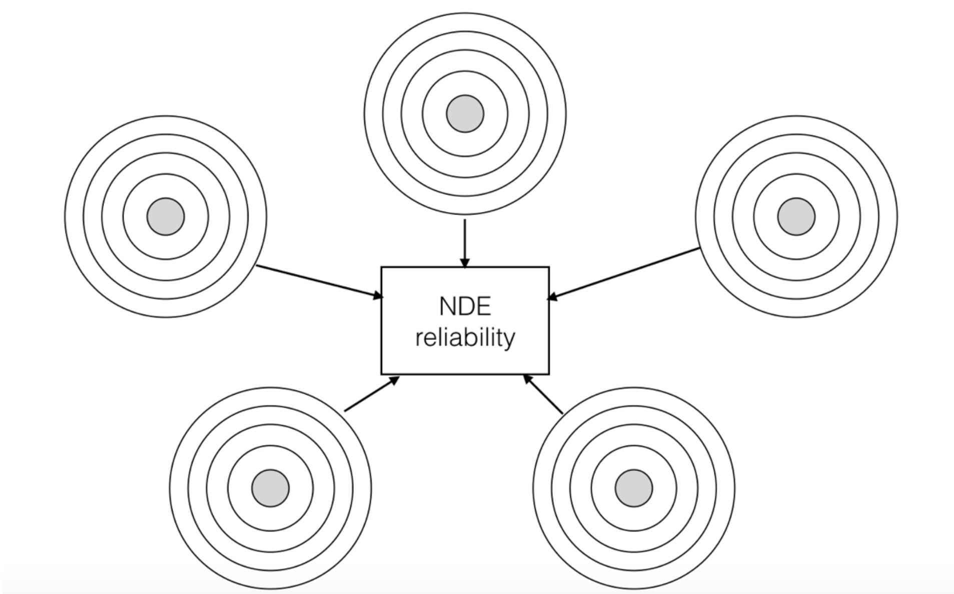


Figure 3-3
Socio-ecological systems model of group performance on NDE reliability where each circle represents hypothetical personnel

4
OPERATING EXPERIENCES AND FIELD NDE OBSERVATIONS

In addition to the research literature, we reviewed actual operating experiences and visited two nuclear power plants in the eastern United States to observe NDE (manual ultrasonic examinations) first-hand. The purpose of the field observations was to create and refine a set of common work artifacts associated with human factors—persona and scenarios. These artifacts are used to describe and better understand the individuals and environments involved in a complex system.

Operating Experiences and Flaw Discovery

Operating experiences in the United States were summarized by the Institute of Nuclear Power Operations (INPO) during a presentation at the 2014 EPRI NDE Technology Week [6].

A 2014 review of NDE reliability by INPO found 16 NDE events since 2007, including four overcalls and one near miss [6]. Of the 11 missed flaws, 8 were found later with ultrasonics, 1 was found with destructive evaluation, and 2 flaws resulted in leaks.

These data suggest an exceptionally low error rate given the frequency of occurrence of examinations. Of course, it is critical to note that the NDE failure rate is unknown—the industry cannot identify 100% of flaws that have been missed in the field; only some percentage of missed flaws have been discovered. Given such an exceptionally low error rate of NDE failures, this implies that there is virtually no obvious way to decrease an already extremely low metric. A similar analogy could be made to the amount of airline crashes in aviation, which are undeniably infrequent. Nonetheless, human factors work in aviation has focused on improving training of pilots, standardization of procedures and documentation, and equipment usability and standardization—all areas of potential future research in NDE.

On-Site Observations

Human factors and EPRI staff traveled to two nuclear power plants to observe manual UT. At one site, the observer participated in pre-evaluation briefings and training, then accompanied examiners on their inspections of feedwater and steam generator UT, while wearing all examiner-required protective equipment. At the second location, observers participated in a pre-job briefing, calibration of a scope device, and one manual ultrasonic evaluation during an outage. Each visit required approximately three days on site; each human factors observer was accompanied by EPRI and utility staff.

There were several positive outcomes of these observations for the current human factors project, including the following:

- Confirmation of assumptions about the stress and cognitive load on examiners during NDE
- Refinement of our understanding of important details impacting manual NDE, including the importance of travel, long hours, seasonal (periodic) nature of the work, the interactive effects of complex briefing information, multiple interruptions, paperwork, cross-referencing, instrument setup, materials preparation, variable weld compositions and complex piping configurations, environment variables, differences in equipment used across sites, and myriad cognitive and physical demands placed on examiners
- Discussion of utility- and site-specific differences in evaluation design and practice
- Increased confidence in the usefulness of human factors paradigms for NDE
- Validated consistency of practices across more than one location

An especially unexpected observation concerned the seasonal nature of NDE and the pressures created simply by transitioning between plants for inspectors; the literature review did not highlight these issues. The travel requirements and differences in plants (for example, administrative controls, site access and badging, and plant design differences like boiling water reactor versus pressurized water reactor) as well as long hours and the pressure of knowing a plant is out of service until work is completed present a cumbersome backdrop to an already stressful role. Further, utilities'/owners' business cultures vary; for example, different utilities have shorter time frames for the same outage scope. The shorter time frames create pressure for all crafts while adhering to all safety regulations imposed by the U.S. Occupational Safety and Health Administration (OSHA). In addition, recall of site-specific information, cross-referencing numerous manuals, and adapting to equipment at each site were surprisingly taxing. At one observation, Level III examiners struggled for approximately 50 minutes with setting up equipment, a notable example of usability challenges that hinder examiners' work. Examiners were continually interrupted as they prepared to conduct NDE, another observation that can be expected to reduce their performance as they frequently shift attention (for example, general National Academy for Nuclear Training eLearning [NANTel] computer-based training [CBT] versus site-specific CBT differs from utility to utility). Finally, spending an entire day of training seated in a classroom or office to review procedures at a new site, followed by substantial time required to don safety equipment—all before even beginning the NDE work—were also surprisingly tiring.

Because human factors work is based on direct observation and measurement of human behavior in context, these site visits mark a critical step in understanding how to prioritize the equipment, tasks, and training for NDE examiners. It was quite clear that even a brief observation of these professionals was an overwhelming and exhausting experience for a layperson. The observations also provide a counterpoint to the literature review: seeing and experiencing NDE in a nuclear environment first-hand validates and corrects many perceptions that researchers might have from simply reading about this activity or hearing second-hand accounts.

Specifically, observations ensure that important system and human performance variables are considered, even when they might receive limited attention from theoretical approaches (for example, one conceptual model omits training, and another omits equipment design). Observation also allows researchers to appropriately prioritize the potential importance of differences between training and field environments, through direct experience of the many variables thought to impact NDE examiners.

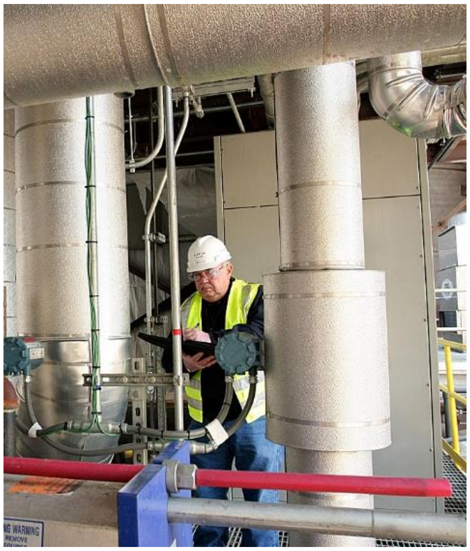
Human Factors Work Product: Persona Design

Personas, or narrative descriptions of critical individuals in a complex system, are used to communicate important human characteristics that might impact the system. In the case of NDE, examiners are typically divided into three subgroups based on their level of certification. The two field observations informed personas for each level of NDE examiner, which includes details about the individual, the job, and the challenges facing him or her (note that most examiners are men, so the personas reflect this fact).

These personas, as shown in Figures 4-1 through 4-3, were developed as composites of information gleaned from the research, then refined with additional details from the field observations. They will serve as generalizations of the individuals and environments encountered and support future research efforts that target individual performance of NDE examiners.

Levels I–III have specific qualifications [7]. Concern has been expressed about the limited time that Level I and Level II examiners can spend with Level IIIs to ensure that knowledge is passed down properly. Further, the industry as a whole has a limited pool of qualified and experienced Level IIIs.

In addition, Level I is a challenging phase in the NDE career. Some utilities have not been able to bring on Level I examiners due to cost factors and because their skills and experience can be limited. Yet the job experience is an opportune way to advance and develop into an effective and knowledgeable Level II.



About the Job

- Usually works 40 hours a week, with longer hours during outages

Spends about a third of his time

- In the field
- In the office
- Providing training and mentoring

Sam

Sam has worked in the nuclear power industry for more than 20 years. He has been Level III certified for 15 of those years. He is proud that his job helps to ensure the integrity and reliability of nuclear power plants. He knows that examiners like him are what protect the public and ensure the safety of the entire industry. As a mentor, Sam wants to pass along his knowledge to those newer to the field.

"*Our vigilance has paid off in identifying potential problems before they become a significant concern."*

About the Challenges

The aging of the units requires even greater attention to increased chances for failures.

Requirements for documentation have increased, and he worries that this takes time away from doing the job.

As he has gotten older, the physical demands at work have become more taxing. He appreciates leaving most of this to younger examiners.

His military and industry background have given him rich experience in the field, and he wonders if he will be able to convey this to newcomers.

Figure 4-1
Level III examiner persona

Adam

Adam has worked in the nuclear power industry for more than 12 years. He has been Level II certified for eight of those years. He prefers to work in the field, so he has postponed pursuing his Level III certification. He's proud to be one of the few with his experience and skills, and works hard to not become complacent—after all, there has never been an undetected flaw on any of his inspections. As a mentor, Adam wants to pass along his tips and tricks to the younger guys, the ones that make him responsible for everyone's safety.

"I like the action during outages and when working in the field—time flies by."

About the Job

- Usually works 60 hours a week or more during outages, and has time off usually in the summer and winter months

Spends about a third of his time

- Review historical data and procedures to review Level I's work
- Calibrating equipment and teaching Level I to do it
- Conducting an inspection while teaching his Level I
- Completing documentation and reporting findings with his Level I

About the Challenges

Pressure to not make mistakes; he does not want to be known as "that guy" who made a mistake because it would go down in history forever.

Must oversee his Level I, double-check and cross-reference all the procedures and applicable documents to ensure that all the paperwork and the calibrations are done correctly while dealing with the industrial environment with health risks, background noise, and activity.

He is the go-to person for both his Level III and his Level I, who both depend on him. His cognitive load is extremely large.

He hopes to have a good relationship with his team members to help work go smoothly—that open and casual communication will help mitigate the stress.

Figure 4-2
Level II examiner persona

About the Job

- Usually works 60 hours a week or more during outages, and has time off usually in the summer and winter months

Spends a fourth of his workday doing each:

- Review of historical data and procedures to fill out UT examination sheets
- Calibrating equipment
- Assisting in the inspection
- Reporting findings with his Level II

Harry

Harry has worked in the nuclear power industry for three years. He has been Level I certified for two of those years. He appreciates working in the nuclear industry because of the challenges of its complexities and technologies as well as the great pay and schedule. Harry wants to learn as much as he can from the Level IIs and IIIs with whom he works. He appreciates having the guidance of an experienced Level II, who gives him lots of hands-on experience.

Harry is a bit of a risk-taker: his hobbies are working on cars and motorcycles, which has honed his mechanical skills, and he likes scuba diving.

"I want to achieve Level II certification as soon as possible so I can do and earn more in my job."

About the Challenges

With the busy pace during outages, he hopes that he learns quickly from his mentors and gets his Level II training paid for by his employer soon.

Must pay careful attention while filling out his paperwork, cross-referencing documents, and calculating numbers mentally. Meanwhile, the workplace is a potential health risk with much background noise and activity.

He knows safety is first but is still getting used to all the regulations and remembering to put his hard hat on, grab his safety glasses, and so on.

Finding the components is still a challenge for him. He relies on his Level II and finding a common area to count down each weld, ensuring he is on the correct one.

He is not used to 80–90°F (27–32°C) working conditions and finds it stressful.

Figure 4-3
Level I examiner persona

Human Factors Work Product: Scenarios

Several scenarios were developed that describe various aspects of an examiner's job. Scenarios were developed based on no flaws, potential flaws, pre-examination requirements, the NDE exam itself, and calibration (see Figures 4-4 through 4-7). These scenarios will inform, guide, and illustrate future human factors research endeavors.

No Flaw	Adam and Harry have spent the morning preparing to inspect a weld during an outage. The manual ultrasound evaluation itself takes about 30 minutes, and they find no indications of flaws. They wrap up the assignment for the day by completing the necessary paperwork.
Potential Flaw	Harry and Adam are conducting a manual ultrasound evaluation of a weld on a large pipe. They believe that they have identified a crack-like indication, and are documenting details in their notes. They plan to follow up with Sam (the Level III) after they leave the work area.

Figure 4-4
Scenarios for manual ultrasound inspection

Pre-Site Preparation (about one to two days, on average)

- Qualified to perform job
  - UT Level II
  - Performance Demonstration Initiative (PDI)
- Personal paperwork for security clearance in advance of visit (PHQs)
- Personal logistics
  - Book travel
  - Know what kind of clothing to wear (check weather)
  - Take the trip
  - Check into hotel
  - Find where to eat
  - Find the plant

On-Site Administrative Preparation (about two to three days, on average, for on-site)

- Security screening of person and vehicle for entry to physical site
- Badging paperwork as appropriate
- Site-specific CBT and performance training
  - Site-specific training on filling out paperwork

Figure 4-5
Before NDE work begins ("First Day in the Life of")

Preparation (about half a workday, on average)

- Receive pre-job brief
- Receive pre-task brief for assignment
- Obtain procedure package, including drawings of plant location and component configuration, historical data, exam area required to meet code
- Review procedure and determine equipment requirements
- Assemble qualified equipment (transducers, cable, and so on)
- Calibrate equipment
- Review and pack up (prepare) equipment and tools, label couplant with chemical label sticker

Suit up in protective clothing if going into radioactive area

Wait for okay to proceed (**Note**: Often, inspectors can wait several hours for others to provide ac power, erect scaffolding, remove insulation, prepare the examination surface, and so on before they can start work. However, the inspectors must be there ready to go to avoid any delays due to their tasks.)

UT Evaluation (about 30–45 minutes, on average, per weld)

- Enter work area
- Access component
- Scan for potential flaws, such as cracks along welds
- Document findings if any on a notepad
- Clean pipe with napkins/squeegee
- Pack up tools and equipment
- Leave work area

Lunch (Note: Once in the plant scanning, the examiner is likely to exit only once the tasks are complete.)

- Return to work area, if needed, to evaluate more welds

Completion (about 2 hours, on average)

- Decontaminate as needed
- Complete documentation and follow up as needed

Figure 4-6
Manual UT high-level tasks ("Day in the Life of")

8:15am	Adam, Level II, begins filling out his UT Calibration Examination sheet in a work trailer. Meanwhile, another Level II is teaching Harry, Level I, how to calibrate his equipment just a couple seats away from Adam.
8:27am	There is a feeling of anxiety in the room as Harry is bouncing his leg up and down and mentions “it’s really cold in here.” Adam acknowledges Harry and looks up at the air-conditioning unit, notices that it is 60°F (16°C), then goes back to his paperwork.
8:43am	Harry is done with his calibration, and Adam moves over to begin calibrating his equipment as Harry and the other Level II exit the trailer, expressing relief from the cold environment. Adam gathers procedures, paper towels, couplant, scope, his UT Calibration Examination sheet on a clipboard, a pen, and a trash bin to discard used paper towels.
8:48am	Adam looks at the serial number on the transducer and asks a Level I to verify the numbers for him.
8:50am	A colleague comes into the trailer to ask Adam a question about the equipment.
8:52am	Adam voices that he is having trouble remembering how to do the Distance Amplitude Correction (DAC). “That’s the thing, you know how to do it and then you don’t do it for six to eight months (or years), and then you have to remember how to use it.”
8:53am	A Level II, the one who was helping Harry, walks in and helps Adam. They both struggle to remember how to do the DAC, and the other Level II was just doing this a few moments ago with Harry.
9:22am	Adam is done calibrating. He puts on his helmet and safety glasses and walks to the next trailer to pick up additional equipment needed for his day (plastic bag to wrap his scope).

Figure 4-7
Manual UT (“A moment in the life of”): calibrations

Conclusions

Important conclusions resulting from the review of operating experiences and on-site observations are as follows:

- In nuclear power plants, flaws in welds are not frequently detected because of good construction and maintenance practices.
- As nuclear power plants age, more flaws can be expected to occur.
- In the dozen or so incidents reported in the past decade, only a handful at most have had significant financial consequences, and none negatively impacted human health or safety but have precipitated industrywide perceptions of problematic NDE reliability.
- Anecdotal evidence, based on conversations with nuclear industry personnel, indicate that most inspections carried out by NDE examiners in the United States have not uncovered flaws. Also, it appears that a considerable number of NDE examiners have never found a flaw in the field throughout their career.
- Based on the number of NDE procedures performed and the number of NDE failures subsequently documented, the rate of identified human error is low, although it might not be an accurate proxy for the unidentified rate of actual error.

Human error cannot be eliminated. However, it behooves researchers, managers, and other stakeholders to better understand the enabling factors that allow NDE examiners to perform effectively despite the number and intensity of stressors that they face; the conditions that maximize and even exceed their coping capacity; and ways to improve their training, preparation, and skill maintenance. Future research must identify ways of alleviating the numerous independent sources of stress to the greatest extent possible. Developing and continuing to refine the targeted personas and scenarios will guide and focus future human factors efforts and support development of additional outcome metrics that better reflect their efforts.

5
RESEARCH GAPS AND OPPORTUNITIES

This review of the existing research literature, operational experiences, and on-site observation provided a comprehensive preliminary examination of both research and practice in nuclear NDE. The critical findings from this effort were as follows:

- The nuclear NDE human performance research literature can be improved with additional up-to-date human factors research on topics such as the following:
  - New opportunities for training and practice available through advances in simulation technology
  - New opportunities for equipment design based on technology advances and human factors guidelines
  - New opportunities to develop and/or implement usability test procedures using simulation
- First-hand examiner experiences and perceptions are rarely considered or measured.
- In the past decade, very few ultrasonic NDE errors in which a qualified procedure performed by qualified personnel failed to correctly implement the procedure have been identified.

Observations of in-field NDE suggest that research has appropriately focused on some of the environmental variables that were observed to be highly impactful to performance, including heat, time pressure, and noise. Nonetheless, the research has tended to overlook how to remediate these variables to reduce the physical and psychological stress on examiners. It is not enough to simply confirm that these variables reduce performance; the critical next step is to identify useful and usable ways of minimizing their impact. Another important avenue of exploration is how existing equipment and methods of reducing the effects of the physical environment could be affecting and potentially increasing the physical or cognitive load on examiners.

Several other variables, such as the pre-job briefing, procedure usability, equipment usability, and adequacy of specific preparation for difficult weld examinations, have not had sufficient research attention. Research has focused on individual difference variables—including cognitive style, attention, and stress response—that are to be expected and might not be directly actionable or could be dependent variables (that is, stress response and coping). Finally, the differences in laboratory training environments and on-site practice have not been carefully articulated, nor their implications thoroughly explored. The area of training might be particularly fruitful for decreasing the gap between the relative calm and quiet of the laboratory and the myriad disruptions and challenges posed by the typical nuclear plant environment. Thus, it seems reasonable to conclude that there is a difference between research and practice.

To identify opportunities for closing the gap between the research and the practice of nuclear NDE, it is critical to focus on the following actionable variables within the traditional scope of human factors practice, in keeping with a socio-ecological model of NDE (shown in Section 3):

- Understanding the differences in training and practice environments between laboratory and in-field NDE, as well as appropriate remediation of the most taxing variables for examiners
- Design and usability of equipment, including readability of displays and documentation, transducer size and shape, and inobtrusive quality of personal safety equipment
- Initial training and ongoing practice of NDE (skill maintenance)
- Specific preparation for an examination, including pre-briefing, procedures, and usability of job aids and equipment
- Leadership and management support for equipment usability, training, and maintenance of examiner skills
- Outage inspections and scheduling constraints
- Examiner interaction and communications with supervisors and utility/client

Combining several of these areas yields a potentially high-impact area for further study—using simulation to train and prepare examiners for NDE, especially for those examinations expected to be particularly difficult, inaccessible, or highly stressful. Simulators have long been a significant part of the training of skilled personnel in multiple industries, including healthcare, military, law enforcement, and aviation.

In control room operation of nuclear power plants, simulation-based training has had significant advantages in reproducing the dangerous, complex, and potentially catastrophic consequences of low-frequency events. Because NDE performed in the power generation industry is like other jobs that demand highly accurate performance under conditions of high stress, simulation has a role to play in the training, pre-job preparation, and skill maintenance of NDE examiners.

For example, the full-size pipe and vessel mockups used for NDE examiners to practice and train on can be expensive and difficult to transport and share, thus reducing their accessibility to the NDE community. Through simulation, virtual data sets can be easily shared and allow more accessible mockups for training purposes.

Furthermore, mockups could be built through the use of lightweight and easily configurable UT scanning simulators to represent field scenarios and could produce a more realistic training environment. EPRI's Virtual NDE v1.0 product is an example of the industry making strides toward addressing this issue [8]: "Currently, Virtual NDE v1.0 is a Windows[1]-based software that allows users to simulate the conditions and functions of manual ultrasonic inspections for training, practice, and testing prior to conducting work in the field" [8]. There are numerous benefits to this technology, and manual UT simulation has the potential for research opportunities as well because it has had virtually no attention by researchers heretofore.

[1] Windows is a registered trademark of Microsoft.

A Human Factors Research Program

As described throughout this report, the field of human factors places humans and their needs at the center of any man-machine system. The first question to be asked in a human factors program is "Who is the user and what are their needs?" Several methods are commonly used to address this question.

As shown in Section 4, this report begins the important process of understanding the user—NDE examiners—by detailing three personas and some aspects of common daily scenarios. This work should continue through a program of research that includes several methodologies that will address the limited direct input from examiners observed in the existing literature. Of these, the most important is task analysis, to better understand the differences between the laboratory and field environments for NDE examiners.

Task Analysis as an NDE Methodology

Task analysis is a process of learning about people by observing them in action to understand how they perform tasks and achieve intended goals. The focus on observation is intended to ground the analysis in real-life task performance and reduce any unintentional biases from second-hand accounts or potentially outdated legacy practices. Task analysis is foundational to understanding any job before research or practical applications are purposeful.

The results of a task analysis are useful for several different purposes and extensions related to staffing and personnel, training, human factors engineering, and safety.

The two best-known methods of task analysis are 1) hierarchical task decomposition (or hierarchical task analysis), in which higher level tasks are broken down into greater detail to support the purpose of the analysis; and 2) cognitive task analysis, in which decision-making, problem-solving, memory, attention, and judgment are a focus.

Several other cognitive methods that could be relevant to the problem of NDE examiner training tasks are the GEMS (generic error modeling system) and its SRK (skills-rules-knowledge) model, which focus on cognitive processes [9]. These frameworks highlight how people with different levels of expertise in handling problems of varying familiarity will lead to different responses. For example, a routine weld examination by a Level III examiner might be handled in an automated, skill-based manner. Examiners with less experience (familiarity) or confidence might use rules to generalize from similar situations or reason from their basic knowledge. These responses can cause different patterns of error. A major goal of any task analysis related to NDE is to ensure optimized training of examiners in preparation for field work. To that end, task analysis techniques as well as other observational and interview methods will help identify potential gaps between current training practices and in-field job task experiences. The work will likely involve observation of training experiences for Level I, II, and III examiners as well as interviews with individuals at each skill level. Although human factors experts have already observed two on-site NDE situations, further observation in the actual work environment will also be required.

Participants in the task analysis will include EPRI NDE trainers and other subject matter experts (SMEs), representative examiners (Levels I, II, and III) from utilities and vendors, and human factors experts. SMEs, trainees, and representative examiners will be the subjects of observation and interview to complete the task analyses. Human factors experts will facilitate inventorying and analyzing tasks and associated needs and provide a hierarchical summary of critical tasks and less critical activities as well as an evaluation of the extent to which training addresses the most critical tasks.

Supplemental Human Factors Methodologies

Table 5-1 outlines a preliminary human factors research program that will begin the work of evaluating and prioritizing issues that most negatively impact field-based NDE as well as those with the greatest potential for remediation. In planning human factors work, a grounded theory approach allows flexibility in addressing the most fruitful avenues of exploration as greater knowledge is acquired. Therefore, the proposed sequence of activities should be viewed as complementary and a potential means of data collection, but not a rigid prescription of activities. Due to the challenge inherent in observing in nuclear plant facilities, it is critical to maintain a flexible approach to gaining direct feedback from examiners. This practical limitation is where other methods could be especially helpful. In addition, collecting information through a variety of different methods allows researchers to triangulate data or arrive at conclusions that are more complex and indicative of reality.

Table 5-1
Research program for NDE applied human factors research

Topic or Purpose of Research	Socio-Ecological Model Level Investigated	Research Questions	Study Design
NDE task analysis	All (especially training)	What are the major skills and tasks associated with each level of NDE examiners? To what extent does task preparation differ in laboratory training as compared with operating needs?	Observation (laboratory and field)
Direct NDE examiner feedback	All	What variables exert the greatest negative impact on performance during in-field NDE? To what extent does training adequately prepare examiners at all levels for in-field work?	One-on-one private and anonymous interviews
Other stakeholder feedback	All	What variables exert the greatest negative impact on performance during in-field NDE? To what extent does training adequately prepare examiners at all levels for in-field work?	Interviews

Table 5-1 (continued)
Research program for NDE applied human factors research

Topic or Purpose of Research	Socio-Ecological Model Level Investigated	Research Questions	Study Design
Equipment design and usability	Equipment design	What differences exist in NDE equipment? What equipment is most commonly used in NDE? Does equipment usability impact NDE examination accuracy and examiner confidence in an exam? If so, how?	Comparative heuristic evaluation and/or observation
Procedure and documentation review	Training and procedures	To what extent does documentation and other training materials use best practices for readability and/or scan ability? To what extent is documentation standardized across nuclear power plants?	Heuristic review

Regardless of methodology, the initial major goals of a user-centered NDE research program based in best practices of human factors are as follows:

- Obtain direct feedback and input from present-day, practicing NDE examiners.
- Identify outcome variables that may supplement and be more sensitive to individual NDE performance than overall NDE reliability.
- Determine the required needs and skills for NDE examiners at the three levels of certification.
- Optimize training, equipment usability, procedures, documentation, and other supports for field-based NDE examiners.

This early proposed work focuses on a variety of evaluative and observational activities that will provide an overview of NDE across all levels of the socio-ecologic model. Additionally, activities that provide an evaluative review of artifacts used in NDE (for example, documentation and equipment) have the potential to be low-risk and high-reward for improving examiner performance, especially for those individuals who travel to multiple plants in an outage season.

6
CONCLUSIONS AND RECOMMENDATIONS

The job of manual UT in commercial nuclear plants represents one of the most extreme working environments possible from a human factors perspective. Examiners perform their jobs under conditions of high physical and psychological stress from myriad sources, including heat, radiation exposure, time pressure, noise, and the possibility of a catastrophic radioactive leak if they fail in their task. These variables induce substantial sensory, physical, and cognitive load on examiners. Therefore, it is important to consider the human factors inherent in the NDE process and how they might be better addressed to improve examiner performance and NDE reliability.

To date, research has acknowledged the stressful nature of the NDE examiner's role. However, several opportunities remain between applied research and actual practice. Today, NDE examiners are subject to extensive preventive measures to improve the safety and reliability of NDE and prevent a catastrophic disaster. Do these efforts help examiners or simply increase their cognitive load and stress? What variables are actionable and result in improved examination ease of use and decreased cognitive load, as perceived by examiners? What steps would improve examiners' perceptions of their own performance? The answers to these questions are unclear at the moment.

Based on this review of the literature, field observation, and review of operating experiences, the discipline of human factors has much to offer to advance the practice of NDE, in terms of both psychological knowledge (for example, stress, performance, and training) and research methodology. By maintaining focus on direct feedback from examiners, improving equipment and procedure usability, training, and using existing psychological knowledge, applied research will focus on the most important goal at hand—decreasing the stress on NDE examiners in one of the most extreme job roles. Common human factors practices—such as on-site observation, task analysis, surveys, and interviews—play a role in increasing the safety, satisfaction, and performance of NDE and in better preparing examiners for the incredible challenges of their difficult but critical role in maintaining public safety.

7
REFERENCES

1. C. Wickens, S. Gordon, and Y. Liu, *An Introduction to Human Factors Engineering*. Addison Wesley Longman, New York, NY, 1997, pp. 384–385.
2. A. Swain and H. Guttmann, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications: Final Report.* Washington, D.C.: U.S. Nuclear Regulatory Commission: August 1983. Report NUREG/CR-1278.
3. J. Spanner, R. Badalamente, W. Rankin, and T. Triggs, *Human Reliability Impact on Inservice Inspection: Phase 1 Summary Report*. U.S. Nuclear Regulatory Commission, Washington, D.C.: March 1986. Report NUREG/CR-4436, Vol. 1.
4. *Human Performance in Nondestructive Inspections and Functional Tests*. EPRI, Palo Alto, CA: 1988. NP-6052.
5. M. Bertovic, B. Fahlbruch, and C. Mueller, "Human Factors Perspective on the Reliability of NDT in Nuclear Applications." *Materials Testin*g Vol. 55, No. 4, pp. 243–253 (2013).
6. C. Larsen. "UT Reliability – INPO Perspective." Presentation at EPRI NDE Technology Week, June 18, 2014. http://membercenter.epri.com; Program 41.04.01; Meetings and Webcasts; 2014 EPRI Nondestructive Evaluation Technology Week.
7. NDT Certification Systems. https://www.asnt.org/MajorSiteSections/NDT-Resource-Center/Codes_and_Standards/NDT_Certification_System.aspx. Retrieved September 22, 2017.
8. *Virtual NDE: Ultrasonic Data Player (VNDE) v1.0.* EPRI, Palo Alto, CA: 2015. 3002005493.
9. J. Reason, *Human Error*. Cambridge University Press, Cambridge, UK. 1990.

A
LITERATURE REVIEW: SUPPORTING DOCUMENTS

Table A-1 provides a list of publications included in the literature review as summarized in Section 2. The unique domain is nuclear power plant NDE with an emphasis on manual UT.

Table A-1
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
1	Ali, Balint, Temple, and Leevers, 2012	Journal	Empirical	In Table 5, conditional probabilities were reported for 12 of 22 PSFs for which data were available or estimated for one test case. 11 of the factors resulted in estimated probabilities of about 0.30. It is not possible to generalize the most important PSFs.	Mathematical model is proposed to allow potential ranking of factors responsible for missing defects or reporting defects when there are none. Technical and, in some cases, human factors data were mined from different industries, including several EPRI reports. One test case was run to exercise the model. "It is not generally possible to guarantee comprehensiveness with this approach. Published data and expert opinion have been used in this instance to guide the choice of errors included, as discussed below." (p. 104)
2	Behravesh, Karimi, and Ford, 1989	Book chapter	Survey	Conscientiousness, stress tolerance, efficient bureaucracy, cooperative, and supportive context	Interviews yielded attributes of what makes a highly competent NDE worker and what makes NDE exams in the field typical, superior, or unusually poor.
3	Bertovic, 2015	Dissertation	Empirical	Degree of automation, team performance	Aim of dissertation is to explore "risks associated with mechanized NDT and find ways of mitigating their effects on the inspection performance."
4	Bertovic, Gaal, Müller, and Fahlbruch, 2009 (reprinted in 2011)	Conference paper	Empirical	Time pressure, mental workload, examiner experience	Based on an experiment in a lab setting, time pressure as measured by a clock seemed to have less influence on manual UT performance than time pressure as perceived by the examiner, that is, mental workload.

Table A-1 (continued)
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
5	Bertovic, Fahlbruch, and Mueller, 2013	Journal	Empirical	Time pressure; mental workload; organizational factors with working schedule, communication, procedures, supervision, and demonstration task noted	This article summarizes five years of research and includes the findings of Bertovic, Gaal, Müller, and Fahlbruch, 2009 (reprinted in 2011).
6	Bertovic, Fahlbruch, Müller, Pitkanen, Ronneteg, Gaal, Kanzler, Ewert, and Schombach, 2012	Conference paper	Empirical	Time pressure; mental workload; “social loafing” (which we may refer to as *team performance*)	This article includes the findings of Bertovic, Gaal, Müller, and Fahlbruch, 2009 (reprinted in 2011). Also, when two examiners conduct an examination, they might do less well than if they work individually because they might over-rely on the other person and/or underperform themselves. Not from article but important to note that the “buddy” system can be effective when used with proper process and training, as done in the military.
7	Bertovic and Ronneteg, 2014	Technical report	Empirical	Procedures	Focus of study was use of procedures in mechanized NDE in disposal of spent nuclear fuel. Important points are that procedures should be developed based on human factors/ usability principles and in conjunction with end users (examiners), and usability-tested with real users (examiners).

Table A-1 (continued)
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
8	Blanchard, Langevin, Spanner, and Thigpen, 2007	Technical report	Empirical	Training	"The higher pass rate of the test group as compared to the control group shows that the structured, hands-on training was effective and should be considered a prerequisite for IGSCC requalification." (p. v)
9	Doctor, Becker, and Selby, 1982	Conference paper	Empirical	Work processes (plotting)	Care and accuracy of plotting axial position of indications appeared to be an indicator of performance effectiveness (preliminary result).
10	Enkvist, Edland, and Svenson, 2001	Technical report	Empirical	Examiner stress/ arousal (time and noise)	"...findings that the performance was higher under the stress condition go against the hypotheses of the present study. However, they go well with the information volunteered by most operators. They reported to feel more challenged under the stress condition" (p. 25). "Operators under stress reduce the demands of the task by simplifying it" (p. 26) and, in this case, it worked well.
11	Harris, 1997	Conference paper	Empirical	Examiner aptitude (a natural ability to do something)	Dynamic Inspection Aptitude Test (DIAT) results were correlated with examiner performance in UT for IGSCC.
12	Harris, 1988	Technical report	Empirical	Procedures/instructions, human-machine interface	Apply human factors principles to instructions/procedures and develop guidelines for operator-control interface design were the two immediately actionable recommendations; nine other recommendations for future research.

Table A-1 (continued)
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
13	Harris and McCloskey, 1990	Technical report	Empirical	Cognitive strategies and decision-making heuristics	Research identified nine cognitive elements and seven UT signal characteristics for successful UT inspection. Most important elements to success: hypothesis testing, avoiding premature foreclosure, using if-then logic, and avoiding disregarding nonsupporting evidence.
14	Lindberg and Selby, 2009	Technical report	Survey	Procedure	The study was a guideline for dissimilar metal weld inspections. Surface preparation is critical; six other good practices, such as procedure training and collaboration between utility and vendor, are encouraged.
15	McGrath, 2008	Technical report	Empirical	Personnel selection, training, procedures, work area preparation (not done by examiners), and organizational culture, including time to complete task	Large effort carried out in the UK. Five recommendations related to the five PSFs are described in more detail on pp. 3–4.
16	McGrath and Carter, 2013	Conference paper	Empirical	Personnel selection, training	This is a subset of McGrath (2008) with a focus on personnel selection and training. Mechanical comprehension and self-awareness to avoid being overly cautious and recognize when a procedure may not apply are the two highlighted aptitude/ personality characteristics.
17	McGrath, Wheeler, and Bainbridge, 2009	Conference paper	Empirical	Procedure	This is a subset of McGrath (2008) with a focus on procedures. Improved performance with the use of a checklist/drawing to support examiners' organized, methodical procedure.

Table A-1 (continued)
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
18	Murgatroyd and Crutzen, 1994	Technical report	Empirical	Work processes (such as calibration, flaw detection, reporting), fatigue, technical skills, shiftwork, training, and simulation	Inspector variability in calibrating equipment and flaws has been demonstrated and quantified, and information has been obtained that will assist in specifying error control procedures (p. 101).
19	Norros, 1998	Technical report	Empirical	Work processes (decision making), supervision (role of foreman)	With respect to decision-making, some examiners (called *inspectors*) focused on personal expertise and skills, whereas others focused on procedures. Also, the role of the foreman was emphasized among examiners as contributing to reliable inspection.
20	Pond, Donohoo, and Harris, 1998	Technical report	Theoretical	Environment (heat, noise), human-machine interface, job design, training, qualification, organizational effectiveness	Study was developed to determine if human factors research yielded information applicable to upgrading requirements in ASME Codes and to suggest future research.
21	Schneider and Bird, 2009	Conference paper	Empirical	Experience, training	Results showed that flaw-sizing capability depends on operator training and experience as well as size of flaw and type of flaw (rough/smooth).
22	Spanner, Badalamente, Rankin, and Triggs, 1986	Technical report	Survey	Task, training, procedures	Two-volume report (note different order of authors for each volume) that documents a literature review and identifies five types of variables that affect UT reliability for in-service inspections. Suggests examiner performance can be improved with application of human factors principles in areas of task, training, and procedures.

Table A-1 (continued)
Summary of literature review

Reference Number	Author(s)	Publication Type	Research Type	PSFs of Interest	Conclusions
23	Weeks, Langevin, and Swain, 2013	Technical report	Empirical	Team performance	The study compared flaw detection capabilities between team and individual scanning. Under particular conditions, equivalent performance was found between team and individual performance.
24	Wheeler, Rankin, Spanner, Badalamente, and Taylor, 1986	Technical report	Empirical	Work processes (flaw detection), human-machine interface	This study focused on the detection of defects using manual UT. Level II and III examiners did not differ in their ability to detect IGSCC flaws. No effect of fatigue was found (although the experimental design was questioned). UT equipment had human engineering deficiencies. It was found that examiner experience with IGSCC did not significantly affect their ability to detect IGSCC flaws.

Note: IGSCC = intergranular stress corrosion cracking

References for Appendix A

1. A-H Ali, D. Balint, A. Temple, and P. Leevers. "The Reliability of Defect Sentencing in Manual Ultrasonic Inspection." *NDT&E International*. Vol. 51, pp. 101–110 (2012).
2. M. Behravesh, S. Karimi, and M. Ford. "Human Factors Affecting the Performance of Inspection Personnel in Nuclear Power Plants." *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 2235–2242 (1989).
3. M. Bertovic. "Human Factors in Nondestructive Testing (NDT): Risks and Challenges of Mechanized NDT." Doctoral dissertation, Technical University, Berlin, Germany (2015).
4. M. Bertovic, M. Gaal, C. Müller, and B. Fahlbruch. "Investigating Human Factors in Manual Ultrasonic Testing: Testing the Human Factor Model." Paper presented at the 4th European-American Workshop on Reliability of NDE, Berlin, Germany (June 2009). (Note *Insight* article often cited is a reprint of this 2009 conference paper: Bertovic et al. "Investigating Human Factors in Manual Ultrasonic Testing: Testing the Human Factors Model." *Insigh*t, Vol. 53, No. 12, pp. 673–676 [2011].)
5. M. Bertovic, B. Fahlbruch, and C. Mueller. "Human Factors Perspective on the Reliability of NDT in Nuclear Applications." *Materials Testin*g Vol. 55, No. 4, pp. 243–253 (2013).
6. M. Bertovic, B. Fahlbruch, C. Müller, J. Pitkanen, U. Ronneteg, M. Gaal, D. Kanzler, U. Ewert, and D. Schombach. "Human Factors Approach to the Acquisition and Evaluation of NDT Data." Paper presented on the 18th World Conference on Nondestructive Testing, Durban, South Africa (April 2012).
7. M. Bertovic and U. Ronnetag. *User-Centered Approach to the Development of NDT Instructions.* Swedish Nuclear Fuel and Waste Management Co., Stockholm, Sweden: May 2014. Report R-14-06.
8. *Nondestructive Evaluation: Improving NDE Examiner Proficiency*. EPRI, Palo Alto, CA: 2007. 1015150.
9. S. Doctor, F. Becker, and G. Selby. *Effectiveness and Reliability of U.S. Inservice Inspection Techniques*. Pacific Northwest Laboratory, Richland, WA: 1982. PNL-SA-10404.
10. J. Enkvist, A. Edland, and O. Svenson. *Effects of Time Pressure and Noise on Non-Destructive Testin*g. Swedish Nuclear Power Inspectorate, Stockholm, Sweden, December 2001. Report 01:48.
11. D. Harris. "Prediction of Inspection Performance with a Dynamic, Computer-Based, Multi-Aptitude Test." *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (Albuquerque, NM, September 1997). pp. 574–578.
12. *Human Performance in Nondestructive Inspections and Functional Tests*. EPRI, Palo Alto, CA: 1988. NP-6052.
13. *Cognitive Correlates of Ultrasonic Inspection Performance*. EPRI, Palo Alto, CA: 1990. NP-6675.
14. *Nondestructive Evaluation: Guideline for Conducting Ultrasonic Examinations of Dissimilar Metal Welds.* EPRI, Palo Alto, CA: 2009. 1018181.
15. B. McGrath. *Program for the Assessment of NDT in Industry*. Serco Assurance, Norwich, UK: 2008.

16. B. McGrath and L. Carter. “Improving Inspection Reliability Through Operator Selection and Training.” Paper presented at 5th European-American Workshop on Reliability of NDE (Berlin, Germany, October 2013).
17. B. McGrath, J. Wheeler, and H. Bainbridge. “PANI and the Role of the Written NDT Procedure.” Paper presented at the 4th European-American Workshop on Reliability of NDE (Berlin, Germany, June 2009).
18. R. Murgatroyd and S. Crutzen. “Human Reliability in Inspection, Final Report on Action 7 in the PISC III Program.” In E. Borloo and P. Lemaitre, P. *Non-destructive Examination Practice and Results: State of the Art and PISC III Results*. Commission of the European Communities, Luxembourg 1994. pp. 99–102.
19. L. Norros. “Human and Organizational Factors in the Reliability of Non-destructive Testing (NDT).” In RATU2: *The Finnish Research Program on the Structural Integrity of Nuclear Power Plants, Synthesis of Achievements 1995–1998*. Technical Research Center of Finland, 1998. pp. 271–280.
20. D. Pond, D. Donohoo, and R. Harris. “An Evaluation of Human Factors Research for Ultrasonic Inservice Inspection.” U.S. Nuclear Regulatory Commission, Washington, D.C.: March 1998. NUREG/CR-6605.
21. C. Schneider and C. Bird. “Reliability of Manually Applied Phased Array Inspection.” Paper presented at the 4th European-American Workshop on Reliability of NDE (Berlin, Germany, June 2009).
22. (i) J. Spanner, R. Badalamente, W. Rankin, and T. Triggs. *Human Reliability Impact on Inservice Inspection: Phase 1 Summary Report*. U.S. Nuclear Regulatory Commission, Washington D.C.: March 1986. NUREG/CR-4436, Vol 1. (ii) T. Triggs, W. Rankin, R. Badalamente, and J. Spanner. *Human Reliability Impact on Inservice Inspection: Review and Analysis of Human Performance in Nondestructive Testing (Emphasizing Ultrasonics)*. U.S. Nuclear Regulatory Commission, Washington D.C., March 1986. NUREG/CR-4436, Vol. 2.
23. *Nondestructive Evaluation: 2013 Team Scanning Assessment Conducted on Behalf of the NDE Integration Committee’s NDE Improvement Focus Group*. EPRI, Palo Alto, CA: 2013. 3002002048.
24. W. Wheeler, W. Rankin, J. Spanner, R. Badalamente, and T. Taylor. *Human Factors Study Conducted in Conjunction with a Mini-Round Robin Assessment of Ultrasonic Technician Performance*. U.S. Nuclear Regulatory Commission, Washington, D.C.: August 1986. NUREG/CR-4600.

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, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI members represent 90% of the electric utility revenue in the United States with international participation in 35 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; and Lenox, Mass.

Together…Shaping the Future of Electricity

3002010462

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