Westinghouse

Westinghouse Electric Company
Nuclear Plant Projects
P.O. Box 355
Pittsburgh, Pennsylvania 15230-0355
USA

U.S. Nuclear Regulatory Commission
Document Control Desk
Washington, D.C. 20555

Direct tel: 412-374-5355
Direct fax: 412-374-5456
e-mail: corletmm@westinghouse.com

Your ref: 52-006
Our ref: DCP/NRC1542

December 2, 2002

SUBJECT: Transmittal of Westinghouse Document, "Man-in-the-Loop Test Plan," WCAP-14396, Rev. 3, Non-Proprietary, dated November 2002

Attached please find WCAP-14396 "Man-in-the-Loop Test Plan" dated November 2002. This report is referenced in the Westinghouse response to NRC RAI 620.008 that has been transmitted to the NRC in Westinghouse letter DCP/NRC1535 dated November 26, 2002. This report is a revision to a previous topical report that was submitted to the NRC in support of AP600 Design Certification.

Please contact me at 412-374-5355 if you have any questions concerning this submittal.

Very truly yours,

Michael M. Corletti

M. M. Corletti
Passive Plant Projects & Development
AP600 & AP1000 Projects

/Attachment
1. WCAP-14396, Rev. 3, "Man-in-the-Loop Test Plan", dated November 2002

D063

Attachment 1

WCAP-14396, Rev. 3

"“‘Man-in-the-Loop Test Plan’”"

dated November 2002

Westinghouse Non-Proprietary Class 3

WCAP-14396
Revision 3

November 2002

Man-in-the-Loop Test Plan Description

Westinghouse

AP1000 DOCUMENT COVER SHEET

TDC: ______ Permanent File: ______ S ______

RFS#: ______ RFS ITEM #: ______

AP1000 DOCUMENT NO.	REVISION NO.		ASSIGNED TO
APP-OCS-T5-001	0	Page 1 of 28	W-J. W. Winters

ALTERNATE DOCUMENT NUMBER: WCAP-14396, Revision 3 WORK BREAKDOWN #: OCS

ORIGINATING ORGANIZATION: Westinghouse Electric Company, LLC

TITLE: **Man-in-the-Loop Test Plan Description**

ATTACHMENTS:	DCP #/REV. INCORPORATED IN THIS DOCUMENT REVISION:
CALCULATION/ANALYSIS REFERENCE:	

ELECTRONIC FILENAME	ELECTRONIC FILE FORMAT	ELECTRONIC FILE DESCRIPTION
6133.doc	Microsoft Word	N/A

ORIGINATOR R. B. Fuld	SIGNATURE/DATE T. P. Hayes / for 11/27/02	
REVIEWERS S. P. Kerch	SIGNATURE/DATE T. P. Hayes / for 11/27/02	
INDEPENDENT VERIFIER T. P. Hayes	SIGNATURE/DATE T. P. Hayes 11/27/02	VERIFICATION METHOD Review
AP1000 RESPONSIBLE MANAGER J. W. Winters	SIGNATURE* J. W. Winters	APPROVAL DATE 11/27/02

*Approval of the responsible manager signifies that document is complete, all required reviews are complete, electronic file is attached and document is released for use.

WCAP-14396
Revision 3

Man-in-the-Loop Test Plan Description

R. B. Fuld
S. P. Kerch
E. Roth
R. Mumaw

November 2002

AP1000 Document: APP-OCS-T5-001

Westinghouse Electric Company LLC
P.O. Box 355
Pittsburgh, PA 15230-0355

TABLE OF CONTENTS

LIST OF TABLES

ACRONYMS

ABWR	advanced boiling water reactor
ANSI	American National Standards Institute
ANS	American Nuclear Society
CBP	computer-based procedures
DCD	Design Control Document
DCS	Distributed (process) Control System
EdF	Electricite de France
EOP	emergency operating procedure
EPRI	Electric Power Research Institute
HFE	human factors engineering
HRP	Halden Reactor Project
HSI	human system interface
HSID	human system interface design
I&C	instrumentation and control
ITAAC	inspections, tests, analyses, and acceptance criteria
KNGR	Korean Next Generation Reactor
MHI	Mitsubishi Heavy Industries
MTPD	Man-in-the-Loop Test Plan Description
NPP	nuclear power plant
NRC	Nuclear Regulatory Commission
PWR	pressurized water reactor
TMI	Three Mile Island
TSC	Technical Support Center
USC	universal soft control
V&V	verification and validation
WPIS	Wall Panel Information System

1 PURPOSE AND SCOPE OF DOCUMENT

This document, the Man-in-the-Loop Test Plan Description (MTPD), describes testing that is planned as part of the AP1000 human system interface design (HSID) process. The MTPD addresses preliminary or "engineering" tests, rather than final or "validation" tests. Engineering tests are performed to support the detailed HSID, described in the AP1000 Design Control Document (DCD), subsection 18.8.1 (Reference 1). In contrast, validation tests are performed to test the acceptability of the final design, per DCD subsection 18.11. Validation tests are addressed by WCAP-15860 (Reference 2).

Engineering tests and associated activities are good engineering practices, which reflect an iterative design process and satisfy applicable regulatory guidance (subsection 8.4.6 of Reference 3). The MTPD provides the following information on engineering tests:

- Defines their purpose
- Describes acceptable testbeds
- Identifies human system interface (HIS) features to be tested and key issues to be addressed
- Describes acceptable performance measures, analysis techniques, and test designs
- Describes test participants needed
- Outlines test procedures to be developed
- Explains how results will be applied

Each topic is detailed in a subsequent section of the MTPD. The MTPD does not provide detailed test procedures for engineering tests. A detailed test procedure will be prepared for each engineering test that is performed.

2 ENGINEERING TESTS

2.1 PURPOSE OF ENGINEERING TESTS

Engineering tests are performed to support the detailed HSID process described in the DCD, subsection 18.8.1 (Reference 1). For AP1000, engineering tests are defined as:

> Preliminary tests whose purpose is to obtain feedback from prototype design products early in the design process.

The use of engineering tests is a good engineering practice, which reflects an iterative design process and satisfies applicable regulatory guidance (subsection 8.4.6 of Reference 3). By providing feedback early, before the detailed design is complete, engineering tests can help to improve the design and to avoid problems in the final product. Engineering tests also may offer concrete insight on questions that cannot be resolved logically (for example, by guidance or analysis). Finally, results from engineering tests provide evidence of design adequacy. Engineering tests thus serve to increase confidence and reduce project risk in the design process.

Engineering tests are performed to obtain empirical results that can be applied directly to understanding and improving the design product. More specifically, engineering tests are designed to produce the following types of results for the prototype design:

- Design-specific operating experience
- Confirmation of necessary performance and integration
- Identification of specific problems
- Subjective feedback from expert users and observers

The methodology of engineering tests, which is descriptive, is distinct from the methodology of hypothesis tests; this is discussed in subsection 2.4.1. Also, engineering tests, which are preliminary, are distinct from validation tests, which are final. Validation tests are performed to test the acceptability of the final design, per DCD subsection 18.11. Validation does not directly support the design process because at this late stage, only serious failures to meet requirements will result in design improvements. In contrast, engineering tests permit specific problems to be addressed before the design is finalized.

2.2 ENGINEERING TESTBED

Engineering tests will be performed using dynamic prototypes of the standard HSI features in the AP1000 main control room. Operator and supervisor consoles will both be provided. This set of prototyped features will include a representative subset of the soft controls, workstation displays, advanced alarms, computer-based procedures, and wall-panel information features being developed for the AP1000. This subset of features will be sufficient for the needs of the engineering tests, test scenarios, and data collection. The HSI features specified for AP1000 are described further in section 2.3 of this report.

The prototype control room will be driven by a high-fidelity simulator that models a Westinghouse or similar pressurized water reactor-based power plant. For these tests, the simulator model does not need to be AP1000-specific. Within the scope of planned test scenarios, the behavior and performance of the testbed will be similar to that required by American National Standards Institute/American Nuclear Society (ANSI/ANS) 3.5-1998 (Reference 4) for full-scope nuclear power plant training simulators. Provisions will be made to record the test subjects' performance and their use of the HSI features.

A suitable prototype and testbed facility that meets the guidance of this section is being developed at the Nuclear Automation 286 site in Monroeville. These facilities will be run and debugged before the scheduled start of testing.

2.3 PRELIMINARY EVALUATIONS

The AP1000 HSI is at a preliminary stage of design. This means that the basic concepts for the HSID features have been selected, but the design details are not developed and integrated. The general features (that is, resources) of the AP1000 HSID include:

- Soft controls
- Wall panel information system
- Computer-based procedures
- Computer-based displays
- Computer-based alarms

These are not revolutionary concepts. In general, digital applications in nuclear power plants have progressed conservatively from nonsafety to safety, from custom to standard, and from independent hardware-based to integrated software-based functions. For the AP1000, Table 2-1 lists projects where, over the last 10 years, similar concepts have progressed successfully in Westinghouse-related projects. Based on this experience, the status and maturity of the AP1000 HSI features that will be subject to engineering tests are discussed further in the following subsections.

Useable implementation of the HSI systems must be demonstrated as part of final validation. To support this commitment, 15 human performance issues have been identified that will be addressed by the HSI design (see Table 2-2). The treatment of these issues is proceeding through experience-based evaluation and testing. Engineering tests will contribute to addressing these issues, with a focus on overall integration, and on key issues identified for various resources. A summary of the issues and of prevalidation evidence to address them is provided in Table 2-2.

2.3.1 Soft Controls

The use of computer-based soft controls is routine outside the nuclear industry. Modern technology has made it practical to replace most dedicated analog devices with functionally equivalent digital devices. Tradeoff assessments typically show that the net advantages of digital devices (specifically soft or virtual ones) outweigh their drawbacks. The design of systems to accommodate soft controls (a related but still different purpose from eliminating hard controls) is thus a technical and commercial necessity.

Soft controls have also been applied in nuclear power plants, though the industry is characteristically conservative and new technology is not quickly embraced. As shown in Table 2-1, Westinghouse has been involved with various applications of soft controls in nuclear power plants. Other power plants such as the advanced boiling water reactor (ABWR) and the recent CANDU reactor units have also successfully applied soft controls (Reference 5). While detailed design questions always must be resolved for any particular implementation, the general concept of soft control has proved acceptable in nuclear power plants around the world.

In AP1000, the basic soft control concept is that of a "universal" soft control device. Universal soft controls (USCs) are virtual control devices residing on a nonsafety platform. USCs are able to send control commands to either safety- or nonsafety-grade systems, processes, and components (Reference 6). When properly supplemented with independent and diverse means to execute plant safety and protective functions, the USC concept has been accepted from an instrumentation and control (I&C) licensing standpoint (Reference 7). To date, however, none of the operating nuclear power plants licensed by the U.S. Nuclear Regulatory Commission (NRC) criteria have implemented USCs.

From a human factors engineering standpoint, the USC concept offers advantages to the operator in terms of convenient access, a consistent interface, and a compact workspace. However, the following validation issues, identified in Table 2-2 under "Controlling Plant State," ask:

- Do the HSI features (for example soft controls) support the operator in performing:
  - Simple, operator-paced control tasks? (Issue 11)
  - Control tasks that require assessment of preconditions, side effects, and post-conditions? (Issue 12)
  - Control tasks that require multiple procedures? (Issue 13)
  - Event-paced control tasks? (Issue 14)
  - Control tasks that require coordination among crew members? (Issue 15)

Issues 11 and 14 are judged as directly pertinent to the soft control implementation. In contrast, Issues 13 and 15 are more a matter of the other features capacity to support the effective use of otherwise adequately designed controls. They are primarily questions of integration across features (including the crew). Issue 12 is not perceived as an issue of performance testing, as much as an issue of adequate availability and suitability of information displayed. Verification of information availability and suitability will be confirmed formally for the final design. So the focus for soft control is on Issues 11 and 14.

Soft control tests during the AP600 program were performed to address Issue 11, and more specific issues for consideration are listed in Table 2-3. The tests performed to address Issue 11 (Reference 8) examined the effects of varying degrees of time lag on control, and compared two alternative HSI designs (sliders and pushbuttons). Contrary to expectation, skilled operators were relatively insensitive to the difference between short (2-second) and long (12-second) lags. This suggests that the expected Distributed (process) Control System (DCS) time delays (within the short 2-second interval) will be acceptable in terms of human performance. As to the alternative designs, both formats were usable for soft control, but sliders turned out to be less desirable due to the added demands they place on visual attention.

In addition to these test results, Issue 11 (operator-paced action) is implicitly covered by Issue 14 (event-paced action). This makes Issue 14 the central human performance evaluation issue for soft controls. Thus, in preparation for detailed design and subsequent validation testing, engineering tests will be designed to address Issue 14 with a focus on: 1) the adequate usability of USCs and 2) their successful integration with other HIS features. This will require building and testing one or more prototype USC designs.

2.3.2 Wall Panel Information System (Large-Area Display)

As shown in Table 2-1, Westinghouse has experience with various implementations of large-area video displays in nuclear power plants. Modern control rooms typically incorporate some form of large-area display for one or more of following reasons:

- Gives a shared view of plant status to the crew (big picture and situation awareness)
- Satisfies need for continuous display of selected data
- Addresses concerns for computer-based control rooms (such as getting lost, tunnel vision, keyhole effect, and added navigation effort)
- More flexible than equivalent hardwired displays
- Favorable hardware cost-benefits
- Reduced maintenance
- Well-received by operators

Normal detailed design questions for large-area displays include the type, size, and position of the display hardware, the information content and format of the display, and what manipulable features (if any) will be provided to the operators. These questions will be answered in part by developing the prototype for, and incorporating results from, engineering tests. Nonetheless, the general concept of large area displays has proved to be acceptable in nuclear power plants around the world.

From the human factors engineering (HFE) standpoint, the utility of the wall panel display concept is not questioned. Rather, concern is focussed on whether its implementation is effective. In particular, the following validation issues, identified in Table 2-2 under "Detecting and Monitoring," ask:

- Do the wall panel information system, and the workstation summary and overview displays support the operator in maintaining an awareness of plant status and system availability without needing to search actively through the workstation displays? (Issue 1)
- Does the wall panel information system support the operator in getting more detail about plant status and system availability by directed search of the workstation functional and physical displays? (Issue 2)

- Do the HSI features support efficient navigation to locate specific information? (Issue 3)

- Do the HSI features support crew awareness of plant conditions? (Issue 4)

Issues 1 and 2 are explicitly defined in terms of the wall panel information system. Issues 1 and 4 are addressed to maintaining awareness of the situation. Issues 2 and 3 concern the operators' access to details, a secondary function of the wall panel information system (that is, cue the operator to look elsewhere.) The main issues for wall panel information system are Issues 1 and 4.

To investigate these issues, two Wall Panel Information System (WPIS) tests were performed for AP600. WPIS Test 1 (Reference 9) was a static test to address the situation awareness of single operators with a functionally organized WPIS. WPIS Test 2 (Reference 10) was a dynamic test to address the situation awareness of crews by comparing alternative WPIS formats. Results from both tests demonstrated that operators could effectively use wall panel displays for situation monitoring and event detection, and that wall panel displays were well-received by operators. Test subjects judged a hybrid approach to WPIS (not strictly physical or functional) as most effective. Results also suggested that, for useful and well-organized information, high display density has more benefits than drawbacks (for example, that the need to navigate among workstation displays is reduced.) Significant operator feedback was obtained, for example, on information that should be added to or removed from the WPIS display.

As a result of these tests, and from the experience gained through other facilities that have used wall panel displays, the risk of providing an effective large area display design has been effectively reduced. Findings from the preceding tests will be reviewed and addressed in future large-area display prototypes. Future engineering tests that include WPIS prototypes will focus on overall integration among HSID features, such as access from WPIS to the workstation display features (Issues 2 and 3).

2.3.3 Computer-Based Procedures

Sophisticated computer-based procedures (CBPs) are an innovation largely pioneered by the nuclear industry. As a result of the post- Three Mile Island (TMI) procedure upgrade programs, extensive effort was made to optimize the format of printed emergency operating procedures (EOPs). However, the strengths of printed media notwithstanding, the inherent limitations of hardcopy were still felt regularly in emergency operations. At the same time, the maturing capabilities of computers promised that CBPs would eventually surpass books as an EOP medium.

Preliminary CBP developments in the 1980s were followed by various fully realized EOP implementations in the 1990s (for example, EdF N4; EPRI EOPTS; HRP COPMA; and Westinghouse COMPRO). Nuclear power plant experience and studies (for example, References 11, 12, 13 and 14) have repeatedly shown that such systems provide effective EOP support and improve the net speed and accuracy of the combined human-machine system.

To date, Westinghouse has supplied two such implementations that are fully licensed and in use within nuclear facilities (Beznau and Temelin). The Westinghouse CBPs add diversity to EOP execution with the computer and the operating crew complementing each other for more accurate, reliable implementation of the procedures. By simultaneously monitoring plant parameters and operator decisions, by bringing all necessary information to one location, and by integrating this information with the decisions and actions being evaluated, the CBP system reduces operators' workloads and improves their understanding of the situation. The CBP system also provides a record of actions taken, pending, and deferred, ensuring that the path taken and the current status of procedure execution are clear. Finally, the CBP allows a more compact workspace while being highly compatible with paper versions of the emergency procedures. Thus, from an HFE standpoint, the implementation of CBPs for emergency operations is a proven advancement in operator support.

A similar implementation is proposed for use in the AP1000 control room. The NRC has experience with computer-based procedures, given its development of review guidance (Reference 15) and the feedback it has received from Beznau (Reference 13). Given that such prior implementations have proven effective and acceptable, the risk of providing an effective CBP design for the AP1000 is low. Nonetheless, many of the issues in Table 2-2 either refer to or imply the use of EOPs:

- Does the integration of (HSI features) support the:
  - Operator in responding to single-fault events? (Issue 7)
  - Operator in interpretation and planning during multiple-fault events? (Issue 8)
  - Crew in interpretation and planning during multiple-fault events? (Issue 9)
  - Crew in interpretation and planning during severe accidents? (Issue 10)
- Do the HSI features support the operator in performing:
  - Control tasks that require assessment of preconditions, side effects, and post-conditions? (Issue 12)
  - Control tasks that require multiple procedures? (Issue 13)
  - Event-paced control tasks? (Issue 14)

Each reference above to events (that is, Issues 7, 8, 9, 10, and 14) implies the use of EOPs, and the implementation of EOPs via CBPs is the primary means to integrate information and controls for the operator during events. Reference to events also implies that CBP tasks are naturally event-paced. Since the preliminary CBP design has already proven adequate in actual use, and since realistic CBP testing requires integrated display and control features, Issues 7, 8, 9, 10, and 14 are viewed as integration questions that warrant no separate testing for CBPs.

Issue 12 is not perceived as an issue of performance testing, as much as an issue of adequate availability and suitability of information provided. Verification of information availability (which has expanded greatly with the use of digital system) and suitability will be confirmed formally for the final design. For the design, this is typically focussed on displays; for procedures, this is primarily an issue of final EOP development. In either case, the relevance to CBP design may be minimal.

Only Issue 13 is directly addressed to procedures, and the preliminary CBP design embodies one approach to Issue 13 that has proven adequate. While the issue is not presumed closed before validation, it is regarded as a sufficiently low risk to preclude the need for separate CBP testing. Therefore, in preparation for detailed design and subsequent validation testing, engineering tests will be designed which address Issue 13 as part of the broader treatment of: 1) the adequate usability of CBPs and 2) their successful integration with other HSI features.

2.3.4 Computer-Based Displays

The use of computer-based displays (sometimes described for AP1000 as workstation displays) is to some extent routine in all industries including nuclear. Human factors questions about computer-based displays are, therefore, not really conceptual, but tend to be more practical matters of verification and validation. Still, as computer-based displays increasingly supplant their conventional counterparts, some concerns linger about the loss of desirable attributes and the addition of new problems.

Examples of Westinghouse experience with typical nuclear applications of computer-based displays are provided in Table 2-1. Given that such prior implementations have proven effective and acceptable, the risk of providing an effective computer-based display design for AP1000 is considered low. Still, most issues in Table 2-2 somehow implicate these displays, because computer-based displays provide the basic context from which other design features are accessed and used.

It is, therefore, desirable to identify which if any of the issues are primarily related to these displays. To avoid repetition, it is noted that subsections 2.3.1 and 2.3.3 identified Issues 7 through 15 as integration matters, except for the unique case of Issue 12. Also, Issue 1, which is defined in terms of minimizing the use of computer-based displays, is seen with Issue 4 in applying mainly to the WPIS (subsection 2.3.2). Finally, Issue 5 is directed to the alarm system (subsection 2.3.5). This leaves the following for consideration:

- Does the wall panel information system support the operator in getting more detail about plant status and system availability by directed search of the workstation functional and physical displays? (Issue 2)

- Do the HSI features support efficient navigation to locate specific information? (Issue 3)

- Does the physical and functional organization of plant information on the workstation displays enhance diagnosis of plant conditions and the planning/selection of recovery paths? (Issue 6)

- Do the HSI features support the operator in performing control tasks that require assessment of preconditions, side effects, and post-conditions? (Issue 12)

These issues can be seen to share a general theme: the effective organization and access of display information. Issue 3, about navigation, applies to displays in a straightforward way. But based on experience, adequate navigation within a modern display system can be refined without dedicated testing. Issue 2 is similar to Issue 3, but emphasizes the finer point of integrated access from the WPIS to lower-level displays. Issue 2 warrants further evaluation.

Issue 6 asks about information organization in terms of problem diagnosis and recovery. This enters the sphere of abnormal and emergency operations. In this light, considering computer-based displays in isolation seems unnatural. Alarms and overview information tend to drive detailed diagnosis (see WPIS

Test 2, Reference 10) and diagnosis (even in the face of unusual problems) is often proceduralized. This suggests that Issue 6 should be addressed as another aspect of integration.

Finally, as stated earlier, Issue 12 is not perceived as an issue of performance testing as much as an issue of adequate availability and suitability of the information displayed. Verification of information availability and suitability will be confirmed formally for the final design. The more general, implied aspect of integration between the soft controls and the supporting computer-based displays should also be refined.

It is important for the detailed computer-based display design to be well organized, accessible, and integrated with other features. To ensure this, the preliminary design relies on experience from successful projects, supplemented by specific evaluations (for example, for Issue 2), which will lead to a prototype. The prototype will serve primarily for event-paced integration testing, but will include measures to assess accessibility and integration of computer-based display information, per Issues 3, 6, and 12.

2.3.5 Computer-Based Alarms

Computer-based alarms are increasingly common in the nuclear industry, most often as a supplement to hardwired alarm tiles. From the early use of plant computers for alarm lists, digital technology has come to offer great power and flexibility for alarm processing. Technological advancement has been accepted more easily for alarms, due at least in part to their status as supplemental aids rather than fundamental indications. Also, advanced alarms expand options, rather than force tradeoffs, in design. The same computer-based system can support conventional tiles, flexible alarm lists, and display-embedded alarm indications, if needed.

As shown in Table 2-1, Westinghouse has been involved with various computer-based alarm applications in nuclear power plants. The basic design issues remain: 1) how to choose alarm parameters and setpoints and 2) how to display alarm information most effectively. In addition, the issues in Table 2-2 are considered to identify those that focus on alarms.

Except for the unique case of Issue 12, Issues 7 through 15 are identified in subsections 2.3.1 and 2.3.3 as integration matters. Issues 1 and 4 apply mainly to the WPIS (subsection 2.3.2) while Issues 2 and 3 apply mainly to navigation in displays (subsection 2.3.4). This leaves the following:

- Does the alarm system convey information in a way that enhances operator awareness and understanding of plant conditions? (Issue 5)

- Does the physical and functional organization of plant information on the workstation displays enhance diagnosis of plant conditions and the planning/selection of recovery paths? (Issue 6)

- Do the HSI features support the operator in performing control tasks that require assessment of preconditions, side effects, and post-conditions? (Issue 12)

Issue 5 is directed to alarms, but is not otherwise clearly different from Issue 4 or general integration issues. Issues 6 and 12, while not directed to alarms, suggest possible roles for alarms, but not in isolation from other design features. Again, an integrated test is indicated. In light of these issues, and of

Westinghouse successful experience with advanced alarm systems, no clear need is identified for alarm-specific engineering tests. Instead, in preparation for detailed design and subsequent validation testing, engineering tests will be designed to address Issue 5 in event-paced conditions with a focus on: 1) the adequate usability of alarms and 2) their successful integration with other HSI features.

2.4 GENERAL TEST PLAN

2.4.1 Methodology

The methodology used to evaluate the HSI prototype involves bringing test subjects representing the target user population (for example, experienced PWR operators) into a controlled performance setting, and systematically observing them as they try to use the prototype to perform representative tasks. The objective is to evaluate the ability of the subjects to adequately perform these tasks, and to document specific problems that arise.

The methods of usability evaluations are typically descriptive with emphasis on user feedback. The importance of user feedback is a principle of usability evaluation, and it makes particular sense when the user population is small and specialized (for example, nuclear power plant [NPP] operators). Descriptive methods and the collection of user feedback both typify control room studies, and stand in contrast to hypothesis tests (that is, those based on inferential statistics and formal experiments). Statistically valid hypothesis tests have a limited role in control room studies for the following reasons:

- Breadth of issues
- Complexity of scenarios
- Requisite large samples
- Cost of repetitions
- Turnaround time

The use of descriptive methods is, therefore, a practical necessity. Fortunately, statistical validity is often not a prerequisite for useful development testing. The descriptive usability methods used by engineering tests are detailed in the remainder of section 2.

2.4.2 Measures and Analysis

The following performance measures will be collected and used as a basis for generating results from the test trials:

- Questionnaires (self-assessment, issue assessments, HSI ratings, and workload)
- Observer checklists (performance milestones and completeness)
- Debriefing (questions, comments, and discussion)
- Time history (speed)
- Action history (actions, usage, errors, and accuracy)
- Plant performance (applicable margins and behavior)
- Video tape

Overall, the measures produce data that are either objective (defined as observable and verifiable), or subjective (defined as inherently meaningful responses from knowledgeable participants). Data that does not fit one of these categories as defined, or that is otherwise not cost-effective to collect and process, will be avoided.

Both objective and subjective data will primarily be treated as descriptive. Based on Westinghouse experience in usability testing, subjective data are the most directly useful as design feedback. For AP1000, subjective measures will be developed for each engineering test to ensure that the following concerns are addressed:

- Treatment of each issue in Table 2-2 that is addressed by the test
- Feedback on all HSID features used in the test, per issues in Table 2-3
- Detection of new problems
- Treatment of test-specific issues, if any

The subjective measures directed to tabled issues will be standardized for integration tests. This will permit additional comparisons and/or trending to be performed between tests when successive iterations of similar tests are performed. Examples of typical measurement tools, data, and analysis techniques used by Westinghouse are contained in References 9 and 10.

Both objective and subjective data, as defined, can be readily summarized with simple descriptive techniques. Use of inferential hypothesis testing for engineering tests will be limited to: 1) planned comparisons deemed necessary or 2) post hoc comparisons of interest. Since the principle concern is to identify problems with the prototype design, unfavorable results will be treated as possible indicators and will be used where feasible to guide portions of the subsequent debriefing. Unfavorable results are defined as any failure of objective measures to conform to valid acceptance criteria, or by subjective results that fall more than one standard deviation below mid-scale.

In data collection, it is easy to assume that if everything else works properly, then reliable data collection will simply occur. In fact, there are many common sources of data loss that can ruin an otherwise acceptable test program. Typical problems include the following:

- Sensing or recording failure
- Undetected bad data
- Data mixup
- Storage loss without backup

By definition, undetected bad data (for example, from a simulator bug) is hard to avoid other than by constant scrutiny. However, most of the other problems can be minimized by systematic checking, formal logging, prompt backup, and so on. These actions should be explicit in the procedures for performing specific tests.

2.4.3 Subjects

In actual operation, the AP1000 main control room and associated HSID features will be used only by highly trained, qualified commercial NPP operating crews. The hypothetical group of all such qualified crew members is referred to here as the "target user population."

The target users can be subdivided on the basis of qualification and experience. For AP1000, two subgroups of interest are referred to here as operators and supervisors. Supervisors by definition have longer experience and higher qualifications[1]. Since test subjects may be required to form crews, selected subjects must have significant NPP operating experience in their assigned role (that is, as operator or supervisor). Representative subjects may include qualified operators/supervisors, former operators/supervisors, or simulator trainers.

Test subjects should be drawn from outside the HSID team. This is not a strong requirement for organizational independence, but it will help to ensure that engineering tests provide as much reasonable challenge to the design as possible.

A key question is the number of subjects to be used in each test (that is, sample size = n). Several authors have examined the mathematical models that underlie descriptive usability evaluations. Plotting the proportion of usability problems detected as a function of number of test participants, the relation can be modeled as a simple Poisson process (Reference 16). In essence, each successive test subject tends to reveal fewer findings. Reference 17 continues in this vein to suggest that five test subjects are typically enough to detect 70 to 90 percent of major usability problems in a prototype. Thus, for a typical test iteration, n = 6 subjects (three crews) will be typical. While specific tests may warrant larger samples, these sample sizes have been shown adequate for usability testing. On the other hand, smaller samples are not normally recommended for engineering tests, as the cost of test development is not justified by the increased risk of incomplete results.

The availability of key personnel is critical for efficient testing. In particular, teams of expert subjects can be difficult to schedule effectively, since:

- Their time is in high demand.
- Blocks of time are difficult to obtain, defend, or revise.
- Your tests may not be their top priority.
- They may not belong to your organization.
- Their participation may require travel.

---

1. The concept of a subject "population" implies a distribution of subject characteristics, such as operating knowledge and ability. A range of these characteristics is desired for engineering tests since weaker operators should provide a stronger test of adequate support, while the stronger operators should provide the most expert feedback. Because the distribution of subject ability levels is not highly critical for engineering tests (as opposed to validation, which should ensure the use of less experienced or otherwise marginal crews), the "range of ability" issue will be addressed simply by meeting sample size requirements and by representing both operators and supervisors in testing.

Nonetheless, both testers and subjects can force schedule changes. To minimize the resulting disruptions, it is important to have contingency plans and effective communications established among participants. For example, subjects could be double-booked, once as principal, once as backup. Approaching test dates would then require the test manager to: 1) confirm the principal subjects' availability and 2) inform the backup subjects that they are not needed at this time. This strategy requires more subjects to be available than are to be tested.

Before testing, subjects will receive a full day of training on the testbed prototype. Training content should be sufficient to prepare for the demands of the planned tests. Obviously, such brief training cannot train subjects to the characteristically high level of fully qualified crew members. As a result, this under-training will tend to weigh against positive test results, through no fault of the design itself. This is typically a conservative error. Reuse of subjects in separate test iterations will tend to eliminate the error, and is permitted, so long as test scenarios are not repeated (so reused subjects cannot perform benchmark scenarios in multiple iterations). Each subject's test experiences should, therefore, be tracked as data, since accumulated experience is expected to bear on other measured results.

2.4.4 Test Design

In supporting the design, and in keeping with the descriptive nature of usability evaluation, the concept of a test iteration is informally defined as follows:

> A set of comparable test scenarios performed on a given prototype.

Several implications are intended by this definition. One is that each iteration of testing produces a meaningful set of results for one prototype. Thus, meaningful testing requires no particular number of iterations.

A second implication is that is it the prototype (that is, the design) that iterates. In contrast, benchmark scenarios (and comparable subjects) are used across a sequence of iterations as a form of experimental control. If prototypes (or scenarios or subjects) can not longer be considered directly comparable, then the new test set is not an iteration of preceding tests.

A third implication is that successive test iterations may vary in size (since scenarios are comparable, and set size has no intrinsic interest). The number of scenarios per iteration is initially small (two benchmarks, plus training), but other scenarios (and/or subjects) may be added to any particular iteration (for example, to make tests more demanding, to increase confidence, and to address particular questions) as suits the interests of the design program.

Engineering tests will use prototypes of one or more AP1000 HSID features, as suits the particular purpose of the test and the level of integration being considered. Two sequences of engineering test are presently planned:

- Soft control prototypes
- Integrated control room prototypes

Test iterations in each sequence require separate test documents, as described in subsection 2.4.8.

2.4.5 Scenarios

The selection of a representative test set is a fundamental problem for testing in general and for HSID validation in particular. Issues about HSID adequacy are easily raised by invoking greater scope and complexity, but this is an endpoint to be reached for validation, rather than the starting point. Representativeness is not initially a problem of engineering tests. Rather, engineering tests seek to be efficient in identifying problems and providing feedback to the design. Only when problems cease to be evident is a more representative (that is, larger) test set indicated. Otherwise, a minimal test set (the benchmarks) is sufficient to advance the design. The concept of test iterations, as defined by this report, allows both the needs of initial efficiency and increasing representativeness to be addressed.

For training, it is proposed that demanding normal and less complicated emergency operations be used. For testing, the use of benchmark scenarios is proposed for efficiency and comparability across test iterations in a sequence. The proposed benchmarks will be two different, highly challenging accident scenarios. The two benchmark scenarios will form the core content of each test iteration. Additional scenarios can be added to any iteration as needed.

Each test scenario needs preliminary exercise or "pilot" testing before it is actually used to collect test data. This is particularly true for benchmark scenarios. Complex scenarios interact with simulator software and operator actions in unpredictable ways. The simulator may malfunction, instructions may be misunderstood, or actions and events may proceed much differently than planned. Scenario failures are inconvenient and potentially costly. Though impossible to avoid entirely, such problems can be reduced by pilot testing the scenario through its expected sequence of events.

2.4.6 Use of Results

Results from the engineering tests will be applied to the following objectives:

- To support the refinement of HSID products
- To identify significant problems in the prototype HSID features, which are to be addressed by subsequent prototypes or the detailed design
- To produce remaining HFE requirements needed as input to detailed design documentation
- To ensure integration among the HSI design features and thereby reduce risk for HFE inspections, tests, analyses, and acceptance criteria (ITAAC)
- To compare alternative design features/configurations on an as-needed basis
- To provide evidence that the issues in Table 2-2 have been, and continue to be, adequately addressed
- To provide evidence that testing is being used in support of the HSI design process, as recommended by NUREG-0711, Rev. 1 (Reference 3)

2.4.7 Planning Guidelines

The development and execution of a typical iteration of specific tests nominally requires 8 weeks from initiation to documented results. This estimate assumes that:

- An appropriate testbed with the to-be-tested HSID features installed is available and operating reliably as described in section 2.2.
- There is an approved, preliminary, iteration-specific test plan as described in subsection 2.4.8.
- All personnel required by the iteration-specific test plan are available to perform their scheduled activities.
- All scheduled processes and activities are pilot-tested before actual use.
- Procedures to avoid data loss are specified and used as described in subsection 2.4.2.

Test sequences and iterations are selected to suit the needs and resources of the AP1000 design process. Test iterations will be planned around AP1000 project milestones and will be subject to modification on the basis of changes to the AP1000 development schedule.

2.4.8 Documentation

Separate engineering test sequences are presently planned for prototypes of soft controls and of the integrated control room. The following AP1000 documents will be produced for each iteration of each test sequence:

- Iteration-specific test plan - A specific test plan will be produced for each test iteration. Test plans will be issued as preliminary revisions before the test and as revision 0 after the test. Preliminary test plans will be used for review, comment, pilot testing, and other preparations for the actual tests. Revision 0 of the test plan will reflect the test as it was finally performed. The test plan document will include the following:

  – Specific test description
  – Description of testbed and prototype HSID features
  – Test staff requirements
  – Subject requirements
  – Indoctrination and training materials
  – Detailed scenario specifications
  – Pilot testing guidelines
  – Actual test procedures
  – Measures/data collection tools
  – Schedule
  – Test log template
  – Directions for training
  – Directions for data collection

- Directions for failed test runs
- Directions for debriefing
- Directions for data processing, analysis, and control

- Iteration-specific test report - A specific test report will be produced for each test iteration as a proprietary Westinghouse document. Test reports will include the following:

  - Test plan summary
  - Data
  - Analyses
  - Results
  - Conclusions

The test plans and reports will be formally disseminated to the AP1000 HSI designers and other applicable members of the AP1000 design team.

Table 2-1 Westinghouse Experience With Applications Similar to AP1000 Human System Interface Design Concepts

Plant	Soft Controls	Wall Panel Information System	Computer-Based Procedures	Computer-Based Displays	Computer-Based Alarms
Operating					
Beznau			X	X	X
Borselle		X		X	
EdF N4	X	X	X	X	X
Ringhals-2	X			X	X
Sizewell				X	X
South Texas				X	X
Temelin	X		X	X	X
Vandellos				X	X
Installed					
OKG	X			X	X
Ringhals-2	X	X	X	X	X
Prototyped					
Beznau		X			
KNGR	X	X	X	X	X
MHI Tomari-3	X	X		X	X

Table 2-2 Human Performance Issues to be Addressed by Human System Interface Design (Based on AP1000 DCD Table 18.8-1)

Operator Activity: Detection and Monitoring		Addressed by:
Issue 1	Do the wall panel information system and the workstation summary and overview displays support the operator in maintaining an awareness of plant status and system availability without needing to search actively through the workstation displays?	- OCS-T2R-100 - Integration tests
Issue 2	Does the wall panel information system support the operator in getting more detail about plant status and system availability by directed search of the workstation functional and physical displays?	- OCS-T2R-120 - Integration tests
Issue 3	Do the HSI features support efficient navigation to locate specific information?	- Design experience - Integration tests
Issue 4	Do the HSI features support crew awareness of plant conditions?	- OCS-T2R-120 - Integration tests
Operator Activity: Interpretation and Planning		
Issue 5	Does the alarm system convey information in a way that enhances operator awareness and understanding of plant conditions?	- OCS-T2R-120 - Design experience - Integration tests
Issue 6	Does the physical and functional organization of plant information on the workstation displays enhance diagnosis of plant conditions and the planning/selection of recovery paths?	- OCS-T2R-120 - Design experience - Integration tests
Issue 7	Does the integration of alarms, wall panel information system, workstation and procedures support the operator in responding to single-fault events?	- Design experience - Integration tests - (Issue 8)
Issue 8	Does the integration of alarms, wall panel information system, workstation and procedures support the operator in interpretation and planning during multiple-fault events?	- Integration tests
Issue 9	Does the integration of alarms, wall panel information system, workstation and procedures support the crew in interpretation and planning during multiple-fault events?	- Integration tests
Issue 10	Does the integration of alarms, wall panel information system, workstation and procedures support the crew in interpretation and planning during severe accidents? (TSC)	(Not applicable to engineering tests)
Operator Activity: Controlling Plant State		
Issue 11	Do the HSI features support the operator in performing simple, operator-paced control tasks?	- OCS-J1-008 - Design experience - Integration tests - (Issue 14)
Issue 12	Do the HSI features support the operator in performing control tasks that require assessment of preconditions, side effects, and post-conditions?	- Design experience - Integration tests

Table 2-2 (cont.) Human Performance Issues to be Addressed by Human System Interface Design (based on AP1000 DCD Table 18.8-1)

Operator Activity: CONTROLLING PLANT STATE		**Addressed by:**
Issue 13	Do the HIS features support the operator in performing control tasks that require multiple procedures? (CBP)	- Design experience - Integration tests
Issue 14	Do the HIS features support the operator in performing event-paced control tasks? (soft control)	- Soft control tests - Integration tests
Issue 15	Do the HIS features support the operator in performing control tasks that require coordination among crew members?	- Integration tests

Table 2-3 Design Issues for Engineering Tests and Measures

	Feature	Issue
1	Soft controls	Discrete control actions
2	Soft controls	Continuous control actions
3	Soft controls	Sequences of actions
4	Soft controls	Simultaneous actions on multiple controls
5	Soft controls	Manual intervention of automated control loops
6	Soft controls	Access and retrieval
7	Soft controls	Read current state, target, and limit values
8	Soft controls	Read available / unavailable actions (interlocks, rate limits, tagouts, etc.)
9	Soft controls	Read responses / feedback
10	Soft controls	Detect anomalies and recover errors
11	Soft controls	Detect and respond to errors
12	Soft controls	Duration and variation of indicated response (lag)
13	Soft controls	Keeping pace (time required versus time available)
14	Soft controls	Visibility of actions to observer
15	Soft controls	Substitutes for equivalent features of conventional main control room
16	Large area display	Includes necessary and sufficient content
17	Large area display	Presents coherent, suitable format
18	Large area display	Addresses different modes
19	Large area display	Useful information is easy / hard to find
20	Large area display	Minimizes workstation access
21	Large area display	Minimizes distractions and uninformative details
22	Large area display	Supports monitoring and change detection
23	Large area display	Supports explanation / understanding of status (e.g., turnover)
24	Large area display	Supports crew coordination
25	Large area display	Read current state, target, and limit values
26	Large area display	Information density too high / too low
27	Large area display	Substitutes for equivalent features of conventional main control room

Table 2-3 (cont.)	**Design Issues for Engineering Tests and Measures**	
	Feature	**Issue**
28	CBPs	Keeps operator informed of plant and procedure status
29	CBPs	Allows use of multiple procedures (placekeeping between...)
30	CBPs	Supports deferred steps (placekeeping within...)
31	CBPs	Permits look back and look ahead in procedure
32	CBPs	Information integrated well / poorly
33	CBPs	Checks operator decisions and flags disagreement before implementation
34	CBPs	Monitors continuously applicable steps
35	CBPs	Gives context and preview
36	CBPs	Easy access to details, contingencies, and support information
37	CBPs	Easy access and integration to required controls
38	CBPs	Flags revised status of completed steps
39	CBPs	Keeping pace (time required versus time available)
40	CBPs	Substitutes for equivalent features of conventional main control room
41	Displays	Navigation is effective
42	Displays	Information is accessible
43	Displays	Information is well-integrated in the displays
44	Displays	Displays are well-integrated with each other
45	Displays	Display format and coding are suitable
46	Displays	Necessary and sufficient display content is available
47	Displays	Information density too high / too low
48	Displays	The number of video display units is adequate
49	Displays	The pointing device is adequate for selection
50	Displays	Displays are well-integrated with other control room design features
51	Displays	Made repeated minor errors
52	Displays	Saw opportunity for serious errors
53	Displays	Substitutes for equivalent features of conventional main control room

Table 2-3 (cont.) Design Issues for Engineering Tests and Measures

	Feature	Issue
54	Alarms	Alerting is effective
55	Alarms	Prioritization is effective
56	Alarms	Alarm symbols and coding are easily read
57	Alarms	Alarm features support searching and sorting
58	Alarms	Alarms were shown in meaningful context
59	Alarms	Alarms are well integrated with other displays
60	Alarms	Low priority and suppressed information remains accessible
61	Alarms	Grouping is effective
62	Alarms	Alarms are distracting
63	Alarms	Alarms were helpful
64	Consoles	Accommodates added personnel
65	Consoles	Supports single operator control
66	Consoles	Sufficient console video display units
67	Consoles	Acceptable arrangement
68	Consoles	Sufficient laydown space

3 REFERENCES

1. AP1000 Design Control Document, Chapter 18, "Human Factors Engineering."

2. WCAP-15860, Rev. 0, "Programmatic Level Description of the AP1000 Human Factors Verification and Validation Plan."

3. NUREG-0711, Rev. 1, "Human Factors Engineering Program Review Model."

4. ANSI/ANS-3.5-1998, "American National Standard Nuclear Power Plant Simulators for Use in Operator Training."

5. APP-OCS-GER-003, Rev. 0, "Design Concept Evaluation and Human Factors Engineering Input for the AP1000 Human Systems Interface"

6. CENPD-396, Appendix 4, Revision 2, "Common Qualified Platform Integrated Solution."

7. Letter dated June 22, 2001, S.A. Richards (NRR) to H.A. Sepp (Westinghouse), on "Safety evaluation for the closeout of several of the common qualified platform category 1 open items related to reports CENPC-396-P, Revision 1, and CE-CES-195, Revision 1 (TAC NO. MB0780)."

8. OCS-J1-008, Rev. 0, "Effects of Control Lag and Interaction Mode on Operators' Use of Soft Control."

9. OCS-T2R-100, Rev. 0, "Wall Panel Information System Test 1 Test Report."

10. OCS-T2R-120, Rev. 0, "Wall Panel Information System Test 2 Test Report."

11. EPRI NP-6846 (1990). "The BWR Emergency Operating Procedures Tracking System (EOPTS): Evaluation by control-room operating crews."

12. NUREG/CR-6398, "Evaluation of the Computerized Procedures Manual II (COPMA II)."

13. NUREG/CR-6749, "Integrating Digital and Conventional Human-System Interfaces: Lessons Learned from a Control Room Modernization Program."

14. Pirus, D. & Chambon, Y. (1997). "The computerized procedures for the French N4 series." IEEE Sixth Annual Human Factors and Power Plants.

15. NUREG/CR-6634, "Computer-based Procedure Systems: Technical Basis and Human Factors Review Guidance."

16. Nielsen, J. & Landauer, T. "A mathematical model of the finding of usability problems." INTERCHI'93 Conference Proceedings, ACM, 1993, 206-213.

17. Virzi, R. A. "Refining the test phase of usability evaluation: How many subjects is enough?" *Human Factors*, 34(4), 1992, 457-468.