

## **VALIDATION - EXPERIENCE AND ONGOING CHALLENGES**

**Eric Davey**  
Crew Systems Solutions  
Box 859, Deep River, Ontario, Canada, K0J 1P0  
davey@crew-ss.com

**Keywords:** Validation, lessons learned, challenges.

### **ABSTRACT**

Validation is an assessment activity that generates evidence to confirm whether the performance of an integrated human-machine system complies with operational objectives. Validation is one of several activities that comprise the standard suite of elements comprising the human performance-related design support activities for a project.

This paper summarizes observations and opinions of the author from his experience in designing and conducting system validations in the CANDU nuclear industry. The paper outlines the validation experience base from which the observations are drawn, summarizes lessons learned, describes limitations with current accepted practice, discusses ongoing challenges, and offers suggestions for improvements in future practice.

### **1. INTRODUCTION**

This paper summarizes observations and the experience of the author concerning validation practice within the Canadian nuclear industry. The paper draws on over 15 years of industry application experience with research, initial design, retrofit, and assessment projects. The willingness of utility, design and regulatory colleagues to share and discuss their project validation experience over the years has been most helpful in developing the author's perspective.

Validation is a design confirmation and assessment activity intended to determine the degree to which the integrated use of equipment, operational aids, and formalized work practices by trained users leads to achievement of performance goals. Validation, as currently practiced, is a culminating activity, generally conducted just prior to, or part of the commissioning of a system for active service. The definition and conduct of validation activities is set out in industry human factors engineering regulatory standards [Canadian Nuclear Safety Commission, 2001; O'Hara, Stubler, Higgins and Brown, 1997) and project specific program plans.

### **2. BACKGROUND**

#### **2.1 System and Human Performance and Validation Scope**

Variability and error in the performance of the operations of technical systems can occur for a number of reasons. Performance below expectations can be the result of system design, configuration, environmental, operational practice, and/or personnel factors. While the influence of many of these factors on overall system performance can be controlled to minimize performance variability or error via design practice, several factors are less controllable and can be ongoing sources of performance variability or error (e.g., operational environmental factors such as alarms, communication, or work requests).

A major contributor to operational performance variability and error are factors associated with individual human performance and their interaction with system and work environment features. For example, human performance is prone to variability and error due to limited attentional resources, individual biases, and modification of rules and models of system operation with time, based on experiential knowledge.

As a consequence, validation of the performance of technical systems should be approached with the recognition that variability in performance and operating errors may present ongoing challenges to consistent achievement of system performance goals. Thus, validation scope should extend beyond limited performance benchmarks to characterizing the effectiveness of individual and integrated use of system features in controlling performance variability and minimizing the adverse impacts of operating errors when they occur.

## **2.2. Validation Experience Basis**

The author's opinions on validation practice discussed in this paper are based on over fifteen years experience with the formal assessment of system performance characteristics for human-machine systems in nuclear power plant applications. The scope of assessment has included measurement of integrated and component system performance, and characterization of improvement opportunities for applications ranging from single user instruments to multiple work teams in high fidelity control room simulator environments. Examples of specific projects include:

- Confirmation of the adequacy of emergency response under minimum staffing conditions for an isotope production reactor facility.
- Confirmation of the adequacy of alarm response practices for a multi-process isotope processing facility.
- Confirmation of the adequacy of crew-system performance for a reactor maintenance tool to drain and refill fuel channels.
- Characterization of the safety adequacy of crew operations associated with radioactive gas accumulation, confinement during decay, and discharge for an isotope processing facility.
- Characterization of operational improvements in operator response to major plant upsets with a new control room annunciation system (Feher, Davey, Rivera and Lupton, 1996).
- Characterization of operational improvements in lead operator and team response for plant maneuvers and outage conditions with a new control room annunciation system (Feher, Davey, Rivera and Lupton, 1996).
- Characterization of operational improvements in single alarm investigation and response with an alarm interrogation display in place of conventional paper-based alarm response guidance (Feher, Davey, Rivera and Lupton, 1996).
- Characterization of improvements in crew safety assessment in major plant upsets using a dedicated safety parameter monitoring display in place of former manual data collection, calculation, and trending (McIntyre, Howard, Davey and Feher, 1998).
- Assessment of the operational benefits of an emergency procedure advisor application in place of conventional paper based procedures.

- Characterization of the operational improvements in shift log preparation, distribution, and use with an electronic shift log application in place of conventional hand-written records.

#### **4. LESSONS LEARNED**

Key lessons learned through this validation experience include:

- Operational Context and Principles - To adequately understand the expectations to be assessed with a system, we have found it important to first understand the operational context (i.e., operating states and strategies) and principles on which the system design and operation is based.
- Identification of Assessment Issues - To acquire a sound understanding of the potential operational issues associated with proposed system use, we have found it useful to employ a diversity of information gathering approaches and sources.
- Importance of Understanding Operator Situation Awareness (SA) - To understand observed system and operator performance, we have found it beneficial to measure operator situation awareness; since operator understanding of plant status in comparison with operational objectives provides the fundamental context in which individual tasks are performed.
- SA Measurement - We use a method of SA measurement that is based on determining individual or team situation understanding from observations of the content of natural verbal communication between team members. This technique permits assessment exercises to proceed to completion without disruptive and unnatural interruptions for questioning of participants.
- Diversity in Assessment Measures - Comparisons of data obtained from multiple assessment measures can highlight important aspects of system and human performance that are not observable using general or single dimensional measures of task success.
- Aligning Expectations - Regular communication with the system owner and regulatory representatives can ensure that there are timely opportunities for concerns with validation benefit, approach, and resource needs to be discussed and resolved.

#### **5. OBSERVATIONS**

As an assessment and design confirmation activity, the emphasis in validation practice is on demonstration of expected performance in an operational setting. Due to the effort and resources required for such integrated tests, budget and schedule constraints often limit the validation scope undertaken. Consequently, validation effort often focuses only on a few of the operational situations likely to be encountered during actual system in-service use, so that system performance is not assessed and confirmed in all operational situations in which the system is to be used.

While the adoption of formal validation practices as part of human factors engineering programs over the past ten years has been a positive step forward, several aspects of current practice limit the validity and application of current validation results, for example:

- Variability in Application - Validation is most rigorously applied to 'identified' system changes under the auspices of human factors program plans. However, many changes in operating practice, procedures, and workplaces with equivalent operational impact are implemented and placed into service with less formal operational testing.
- Performance Measures and Criteria - Reliance on general or global measures of task success, and subjective or opinion-based assessment, rather than very task-specific and objective measures and criteria.
- Situational Realism - Validations are seldom undertaken in environments that truly mimic those experienced in operational practice, thus discounting the validity and transferability of results (e.g., task performance in isolation of normal workplace burdens and distracters, or use of optimally trained staff).
- Error and Upset Prevention Effectiveness - Little emphasis on characterizing and confirming the effectiveness and margins provided by human performance related defensive design and operational features.

## 5. CURRENT CHALLENGES

Three additional aspects of design confirmation activities are presenting new challenges to industry validation practitioners, and system design, operations, and regulatory staff:

- Improving Benefit versus Costs

Current validation practice is based on a model employing selective, integrated systems performance tests, apart from, or during system commissioning. This approach offers narrow operational coverage, limited opportunities for system refinement based on validation findings, and can be costly.

We have been exploring alternative models for accumulation of system performance confirmation evidence that offer greater benefit versus cost potential. Two models have been outlined to date:

- Incremental Validation - This approach partitions system performance confirmation activities into smaller assessments and shifts the schedule of assessments back into the design phase. With this approach, performance confirmation evidence is assembled from multiple small assessments. This approach also affords earlier identification of opportunities for system improvement and implementation prior to freezing system design at or prior to commissioning.
- Periodic Observations During Commissioning and Training - This approach relies on accumulation of performance evidence from observations of system and crew performance during planned system operation in commissioning tests and crew training. This approach offers collection of a much broader and richer base of performance confirmation evidence at potentially cheaper cost than defining and staging validation specific exercises.
- Judging the Sufficiency of Defensive Features

Designers and operations staff employ multiple equipment and operational features to promote correct human performance and minimize the occurrence of errors. Examples of these features that are used singly or in combination include:

- Labeling to provide unambiguous information and device identification to promote location and correct selection.
- Visual coding to promote correct information or control identification and selection.
- Grouping of information and controls according to system or function affiliation to promote visual location and correct selection.
- Warnings and Cautions to alert users to the consequences and side effects of operating actions or the consequences of operating errors before actions take place.
- Procedures and written guidance to specify permitted operating actions and action sequences.
- Self-check operating practices for 'on-task' confirmation of information or device control selections or actions.
- Affordances to assist users in recognizing device control possibilities.
- Constraints to limit action possibilities to acceptable choices and ranges.
- Independent verification of information or device control selections or actions to provide enhanced assurance of correct operating choice.
- Undo capabilities to permit users to reverse operating actions and recover from errors.
- Interlocks that physically prevent user actions that may result in unsafe or uneconomic consequences.
- Indications of feedback from process parameters or equipment states resulting from operating actions to provide confirmation of successful operating action success.

The error prevention effectiveness of many of these features can be situationally dependent, being a function of the type of task action, error possibilities and characteristics of the operating environment. Thus, an understanding of these three factors is required to determine the specific error prevention effectiveness of an individual interface or operating feature.

Generally, such level of information detail is not readily available and the human error prevention effectiveness of individual or combinations of features is only known in a qualitative sense. Thus, judgments of the sufficiency of error prevention features can not be made with specific assurance.

- Confirming Results Validity with Operational Change

During operations, plant systems are subject to a variety of operational changes, for example utilities are continually introducing changes to their operational environment. Consequently, the conditions of system operation will begin to depart from the conditions in which system performance was initially confirmed during validation. In such circumstances, a point may be reached where the initial system performance

validation is no longer relevant. Consequently, the system performance may require re-validation to justify continued in-service use.

To address this challenge, designers, operations staff and regulators need to begin identifying the factors that are most critical to maintaining validation relevance for projects that they undertake. Likewise, the merits of alternative ways for characterizing and/or maintaining the operational relevance of system validation results should be explored. Solutions could include:

- Envelope Definition - This involves the adoption of validation approaches that confirm operational performance within a defined envelope of operating conditions rather than for bounding or single scenario situations. As long as key in-service operating conditions remain within the envelope of validated performance, the original validation of the system performance would remain relevant.
- Periodic Re-validation - This approach is dependent on repeated confirmations of system performance across the system's service lifetime. Either time-based or operating envelope excursion criteria could be used to provide the basis for re-validation frequency.
- Ongoing Confirmation - Continuous confirmation of system performance against actual in-service operating conditions would provide a means to maintain validation relevance as operating conditions change. In this case, collection and interpretation of operating data could be used to continuously confirm system performance so that the validation results would remain up-to-date.

## **5. SUGGESTIONS FOR IMPROVEMENT**

To improve the effectiveness of industry validation practice, the following suggestions are offered:

- Alternative Validation Paradigms - Alternatives to the current assessment focus on selective pre-operations, integrated systems testing should be encouraged to explore opportunities for increased assessment coverage, earlier identification of system improvements, and realizing improved benefits versus costs.
- Characterization of Human Error Defenses - Industry organizations should be encouraged to collaborate on and contribute to the development of a defense characterization matrix. Such a matrix would enable comparison of effectiveness results based on single and combinations of features from participating organizations. These combined results could then serve as design guidance and practical assessment criteria for future projects.
- Broadening Assessment Emphasis - The current assessment emphasis should be moved beyond simple operational confirmation to encompass confirmation of design and operational defense effectiveness. Such an assessment extension would provide evidence for assurance of the effectiveness of human error defenses on each application project.
- Characterization of Margins - Establishing and maintaining operating margins provides assurance that operating conditions remain distanced from operational or safety challenges. Where margins are employed, confirmation of actual operating values with design intent should be characterized during validation.

- Ongoing Performance Assurance - The most comprehensive demonstration of the degree to which equipment features, operating practices, and staff skills support achievement of operational objectives is through routine tracking of operational performance. Tracking system performance during operations would provide assurance the system remains capable of achieving operational objectives in spite of ongoing operational change. This would require a shift in validation emphasis from pre-operations testing to encompass on-going operational experience tracking.

To re-confirm error prevention, detection and recovery effectiveness using routine operating experience would require collection of information on the types of errors that occur, their frequency and the degree of feature effectiveness in mitigating their challenge.

The examination of recordable operating events representative of major system performance challenges and breakdowns would not be sufficient for demonstrating the effectiveness of human error defenses. While the occurrence of events offer confirmation of instances of performance failure, the absence of events can not confirm the effectiveness of human error defenses without an understanding of the type and frequency of human error initiated challenges.

## **5. CONCLUSIONS**

This paper has shared the author's insights and opinions regarding validation practice from fifteen years of application in the CANDU nuclear industry. The paper has identified lessons learned, discussed limitations with current practices, identified three challenges to broadening and improving current practice, and offered suggestions for future improvement. Validation in its current or enhanced form is expected to remain a key tool for demonstrating design confirmation in the nuclear industry.

## **ACKNOWLEDGMENTS**

The author wishes to thank CANDU utility, design and regulatory staff who have shared their experiences, insights and concerns with respect to industry validation practice over the past fifteen years.

## **REFERENCES**

- Canadian Nuclear Safety Commission, 2001. Guide to Human Factors Verification and Validation Plans. Canadian Nuclear Safety Commission draft Regulatory Guide C-278, Ottawa, Ontario.
- Feher, M., Davey, E., Rivera, D., and Lupton, L., 1996. Validation of the Computerized Annunciation Message List System (CAMLs). Proceedings of the International Atomic Energy Agency Specialist Meeting on 'Experience and Improvements in Advanced Alarm Annunciation Systems in Nuclear Power Plants', Chalk River, Ontario.
- McIntyre, C., Howard, S., Davey, E., Feher, M., 1998. Operational Assessment of Critical Safety Parameter Monitoring - Findings and Lessons Learned. Proceedings of the Canadian Nuclear Society conference, Toronto, Ontario.
- O'Hara, J., Stubler, Higgins, W. J. and Brown, W., 1997. Integrated System Validation: Methodology and Review Criteria. United States Regulatory Commission report NUREG/CR-6393, Washington, District of Columbia.

