Emerging Technologies and Future of Work, Vol. 117, 2023, 29–38
https://doi.org/10.54941/ahfe1004402

HUNTER Procedure Performance Predictor: Supporting New Procedure Development With a Dynamic Human Reliability Analysis Method

Ronald Boring[1], Thomas Ulrich[1], Roger Lew[2], and Jooyoung Park[1]

[1]Human Factors and Reliability, Idaho National Laboratory, Idaho Falls, ID 83415, USA
[2]University of Idaho, Moscow, ID 83844, USA

ABSTRACT

The Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) is a streamlined software framework for dynamic human reliability analysis (HRA). HUNTER simulates a reactor operator as a digital human twin, providing a platform by which to model human interactions with a digital hardware twin in the form of a simulated nuclear power plant. HUNTER gives realistic insights into human errors, actions, and time frames. Recent discussions with stakeholders in the U.S. nuclear sector have highlighted potential uses for HUNTER to support development of operating procedures for advanced control rooms. As nuclear power plants transition from analog to digital control rooms or develop advanced control systems for new reactors, a unique challenge arises concerning operating procedures. Established procedures for existing plants have undergone multiple iterations, but with the advent of digital control systems in control rooms, there's often a lack of operating experience to shape these new procedures. Such Version Null procedures are a pressing concern for those drafting them and ensuring plant safety. Expanding on HUNTER's procedural capabilities, a specialized version named the Procedure Performance Predictor (P3) is under development. HUNTER-P3 enables those writing procedures to draft new ones and then test them in a simulation to gauge both operator and plant responses. HUNTER-P3 identifies potential operator and procedure level shortcomings, offering a novel way to validate procedures.

Keywords: Dynamic human reliability analysis, HUNTER, Nuclear power, Operating procedure

REVIEW OF HUNTER

The Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER; Boring et al., 2022) is a dynamic human reliability analysis (HRA) tool designed to be simple to use. Initially based on an effort to create a dynamic implementation of the Standard Plant Analysis Risk-Human (SPAR-H; Gertman et al., 2005) HRA method, HUNTER grew to become a standalone software package that allows analysts to use procedures and a linked nuclear power plant model to create a realistic simulation of human performance that can be considered a virtual operator.

The basic structure of HUNTER includes three functional modules:

- *Task module*—which is driven by plant operating procedures
- *Individual module*—which is those factors, specifically performance shaping factors (PSFs), that affect the operator
- *Environment module*—which is a model of the virtual world of the simulation, typically a simulator.

The software implementation of HUNTER includes additional modules necessary to execute HUNTER as standalone software. These include software modules such as a scheduler, which interfaces between the task, individual, and environment, and coordinates Monte Carlo runs to produce distributions of performance outcomes.

Recent versions of HUNTER (Lew et al., 2022) include the use of the Rancor Microworld Simulator (Rancor; Ulrich et al., 2017), a simplified pressurized water reactor simulator that has been used in a variety of studies with student and licensed reactor operators (e.g., Park et al., 2023). The advantages of Rancor center on its simplicity, which allows it to be more readily used than a full-scope and full-scale simulator for studies to collect operator-in-the-loop data, and which features a reduced number of parameters compared to full-scope training simulators. In other words, Rancor is easier to interface with HUNTER than conventional simulators for proofs of concept while also allowing ready collection of empirical data to validate HRA models.

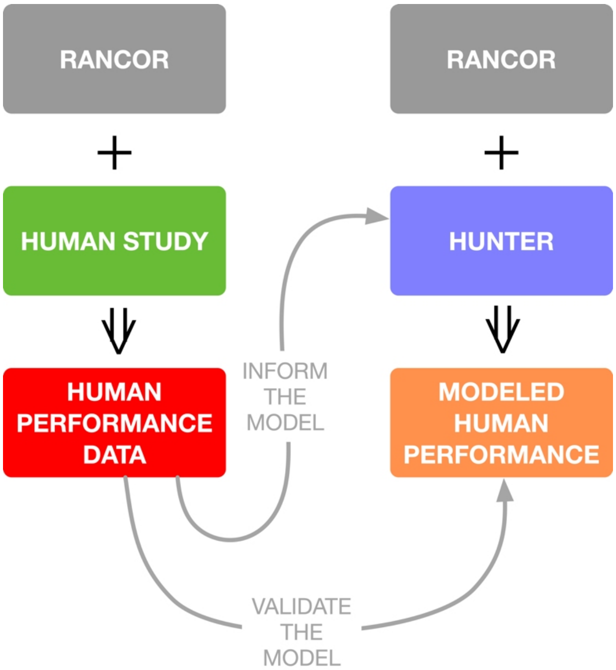

Figure 1: The relationship between HUNTER and simulators for dynamic HRA modeling.

The basic framework for connecting HUNTER with a simulator is shown in Figure 1. Scenarios are run using the simulator and representative operators to collect initial human performance data. The simulator is then coupled to HUNTER, and the human performance data are used to help refine the HUNTER model. For example, when observed in a study, procedural sticking points or timing data can be used to refine and calibrate the basic modeling parameters in HUNTER. A new set of scenarios is then simulated in HUNTER, and simulation outputs can again be compared to available human performance data from simulator runs. If human performance data are collected from two scenarios, e.g., steam generator tube rupture (SGTR) and loss of feedwater (LOFW), one set of data like LOFW would be used to calibrate HUNTER, while HUNTER would then predict performance for SGTR. The predicted and actual SGTR performance data can then be compared for validation of the modeling.

Although preliminary work has focused on using the simplified Rancor Microworld Simulator with HUNTER, the general approach is readily extensible to full-scope plant training simulators. In fact, the Rancor human-system interfaces are based on tools used for prototyping digital upgrades in plant simulators (Boring, Lew, & Ulrich, 2017), meaning Rancor mimics the functionality and advanced programming interface (API) used on full-scope simulators. As such, scaling from the generic Rancor simulator to a plant-specific simulator is a straightforward task. The process depicted in Figure 1 may be replicated for plant simulators assuming human operational performance data are available from the plant simulator to calibrate and validate the HUNTER simulation.

INTRODUCING HUNTER-P3

Much has been written about control room upgrades and the transition from analog to digital systems (e.g., Boring et al., 2019), but relatively little research has been conducted specifically on procedure use with these new systems. An exception is the case of computer-based procedures, where procedures represent one of the technological systems being introduced into the modernized control room (Lew, Boring, and Ulrich, 2018). Despite the minimal research specifically on procedure use amid changing concepts of operations, the procedures used to operate any system of the plant are an important part of the overall human-system interface at the plant.

In recent industry forums to discuss uses of HUNTER, a strong use case has emerged outside traditional applications of HRA for risk assessment. Given the focus in HUNTER on running procedures with a plant simulator, there is a unique and much-needed application of HUNTER to evaluate new procedures. Existing operating procedures at plants benefit from extensive operating experience, industry benchmarking, sharing lessons learned such as through the Pressurized Water Reactor Owners Group (PWROG), and continuous improvement through procedure revisions. However, two new situations challenge this process:

- Plant upgrades that introduce new digital systems in the main control room that require new or extensively modified procedures

- New plants that feature entirely neoteric main control rooms that likewise require new procedures.

These Version Null procedures present potential safety and efficiency concerns for operator performance.

To address this challenge, HUNTER is incorporating a new function called the Procedure Performance Predictor (P3). HUNTER-P3 uses HUNTER's built-in Monte Carlo tools with human performance variability to identify where in procedures there might be error traps. In this manner, HUNTER-P3 can be used to flag problems with the procedures themselves or issues with the execution of the procedures by reactor operators. HUNTER-P3 can serve as a screening tool for novel procedures to help iterate and refine them prior to deployment. Identified error traps serve to prioritize scenarios where empirical evaluation is warranted.

HUNTER includes a procedure authoring system that makes it easy to input procedures to drive the Task Module. A prototype tool called HUNTER-Gatherer uses natural language processing to automate the process of inputting procedures from existing libraries. In this manner, it is possible to use HUNTER-P3 in conjunction with other procedure authoring tools to simulate Version Null procedure performance.

THE IMPORTANCE OF SIMULATOR COUPLING

To realize HUNTER-P3, HUNTER must be coupled to a simulator, as noted earlier in this paper. Boring et al. (2023) explain the importance of synchronous vs. asynchronous coupling for realistic modeling of human-system interactions. Coupling refers to the link between the virtual operator (i.e., HUNTER) and the environment model (i.e., a plant simulator). Asynchronous model coupling occurs with a model code (e.g., a thermo-hydraulic simulation) that is designed to operate without evolving inputs. Asynchronous models take all inputs at the beginning and then run in a batch mode to a defined stop point. For example, an SGTR scenario run might feature a sequence of events as follows:

- *Initial Condition:* Normal operations at 100% power for the starting point,
- *Fault:* Introduction of the fault in the form of the rupture,
- *Mitigative Actions:* Manual responses to the rupture like reactor trip and safety injection, and
- *Termination:* Completion of the scenario at a specified time or upon achieving cooldown status.

This sequence is repeated in Monte Carlo fashion, with slight systematic variations like different times for performing mitigative actions and stochasticity such as normal variability in plant conditions and timings, which produces a range of outcomes for parameters like total leak time or volume.

The key distinction between synchronous and asynchronous coupling is how mitigative actions are performed. In asynchronous coupling (see Figure 2), the mitigative actions are predefined to be performed at specific points in the scenario runs. In contrast, in synchronous coupling, mitigative actions are part of an iterative feedback loop between plant and operator, whereby the operator responds to plant conditions, changes those conditions, and then the plant proceeds from that changed state. The value of synchronous coupling is predicated on three assumptions:

1. Human actions are responsive to emerging plant conditions and cannot be completely determined a priori.
2. Human actions will change plant conditions in a way that meaningfully alters the course of the scenario in an evolving manner that also cannot be completely determined a priori.
3. Timing of human actions as well as selection of specific actions among many possible mitigations changes plant conditions in ways that cannot be fully determined a priori.

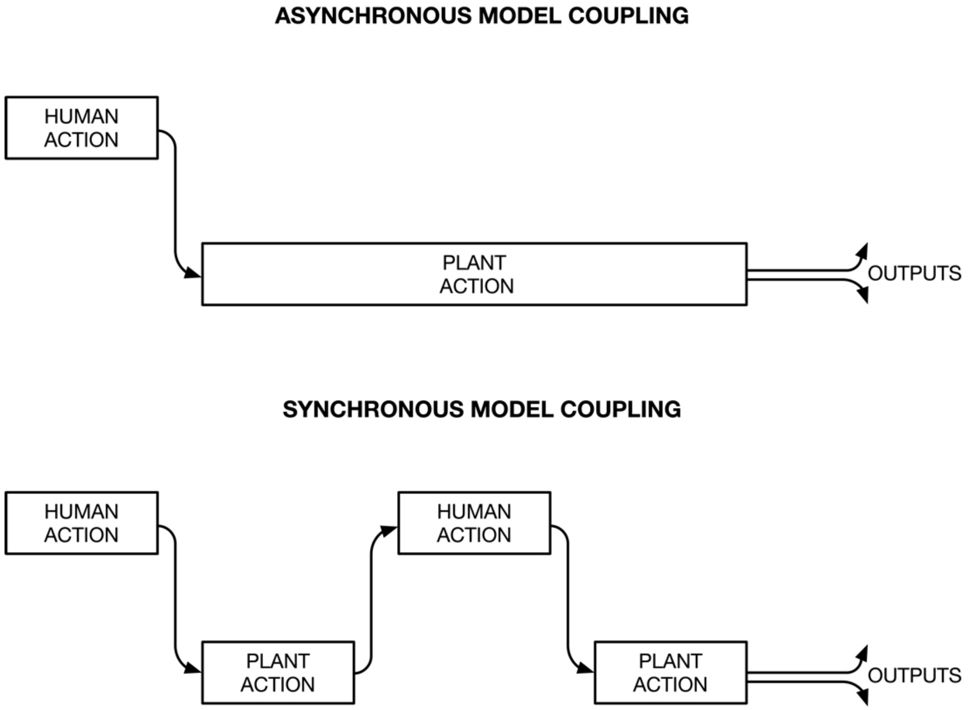

Figure 2: Human-plant interaction for asynchronous and synchronous coupling (from Boring et al., 2023).

The common theme for these assumptions is that there is a myriad of possible plant and human outcomes as the scenario unfolds. For example, a 30-second delay in responding to an upset may change the course of that event in a way that requires completely different actions. Asynchronous models typically do not fully model deviations from the nominal path that may result from the dynamics of the evolving operational context.

Another perspective on asynchronous vs. synchronous coupling is to consider it in the context of normative vs. descriptive models. Bell, Raiffa, and Tversky (1988) delineate normative models as those that predict an ideal outcome, whereas descriptive models reflect actual performance. Asynchronous modeling typically results in normative outcomes—the expected normal or ideal case. Synchronous modeling results in descriptive models—the actual case. The latter is essential for understanding the realistic course of operator behavior when using procedures. HUNTER-P3 synchronously couples its virtual operator representation with the plant full-scope simulator to predict how operators would actually perform when using procedures.

It should be noted that HUNTER-P3 possesses all features necessary to automate plant operations if coupled to an actual plant instead of a simulator. However, the performance of HUNTER-P3 would not fit the normative performance expected of automation. Rather, HUNTER-P3 would provide an operating context that incorporates operator shortcomings. Such decidedly and occasionally imperfect human operation is, of course, seldom the goal of automation.

EVALUATING NEW PROCEDURES

Here we consider the process of how a digital upgrade affects procedures. For example, a newly modernized digital turbine control system (TCS) may largely mimic the functionality of the existing electro-hydraulic control (EHC) it replaced at the plant. However, the new digital interface and control system require slightly different actions by the operator, necessitating new procedure steps or even whole new procedures.

A thorough operating experience review (International Atomic Energy Agency, 2018) can identify potential problem areas with the new system and the use of procedures. A limitation of this approach is that all existing experience may be based on the legacy EHC, and there may be limited performance with novel systems. Moreover, using the example of the TCS, the TCS is one of the first wholly digital control subsystems installed as an upgrade at most U.S. nuclear power plants due to its potential high return on investment through possible power uprates. As the first-of-a-kind installation of a digital system, there is little operating experience to draw upon to ensure the procedures adequately support operator use of the new system.

In the absence of adequate operating experience to provide confidence in novel procedures, the next course of action is to perform empirical evaluations with operators in the loop. This approach is identical to the types of human factors validation activities performed as part of upgrades. Scenarios to represent the range of activities performed with the system or procedure are identified, with a particular emphasis on any critical safety functions. These scenarios capture the continuum from frequent and normal activities to rare and abnormal events. In a TCS, this would cover startup, shutdown, and power evolutions at the plant for normal operations and upset conditions like failed governor valves or grid disturbances. Operators perform these scenarios using the new system and accompanying new procedures, and any performance deficiencies such as confusions, erroneous actions, or response delays are documented and corrected in the system and procedure. This approach is very effective, but it is costly in terms of staffing effort to setup and carry out the studies. Additional challenges are that it is only as effective inasmuch as the scenarios anticipate the actual range of use, and the sample size of operators may be limited, depending on the plant's ability to deviate from operating and training schedules to support engineering and evaluation activities.

HUNTER-P3 presents a novel third approach to identifying issues with new procedures. HUNTER is coupled to the plant's training simulator with the updated control system via the simulator's available API. The API allows monitoring and controlling all simulated plant parameters. Thus, the plant indicators that should be monitored by the virtual operator can be fed into HUNTER, and any control actions taken by the virtual operator in HUNTER can be input into the simulator, allowing HUNTER to function like an actual operator at the control panels. Typically, the API also allows control of instructor station functions, meaning it is possible to start and stop the simulator and insert faults. This functionality is used for the Monte Carlo repeated trials. In this manner, HUNTER-P3 controls the plant's new control system by following the new procedures embedded in HUNTER's individual module.

As noted, HUNTER-P3 simulates the proceduralized activities not in a normative or idealized manner but in a manner that incorporates realistic fallibility of the operators. The individual module accounts for those factors that may impinge on optimal performance. For example, the presence of the PSF for elevated stress may decrease the time to complete the task, while the presence of complexity through multiple simultaneous tasks may increase the likelihood of skipping a procedure step. Hollnagel (2017) suggests there is often a disparity between work as imagined (WaI) and work as done (WaD). HUNTER-P3 captures this at two levels of analysis:

- *Operator level:* The operator does not perform the procedures perfectly due to contextual factors that hinder perfect procedure following, potentially resulting in less-than-perfect plant performance
- *Procedure level:* The procedure does not adequately cover the use context, such that even following the procedures perfectly will not result in perfect plant performance.

Operator level issues in procedure performance result from systematic decrements in WaD that can be accounted through HUNTER in the individual module. Procedure level issues result from WaI not adequately covering the operating envelope of the system. HUNTER-P3 can account for WaI through modeling what-if plant contexts in the environment module, e.g., inserting faults to stress-test the procedure. To avoid confounds between operator and procedure level issues, individual module and environment module factors can be manipulated separately.

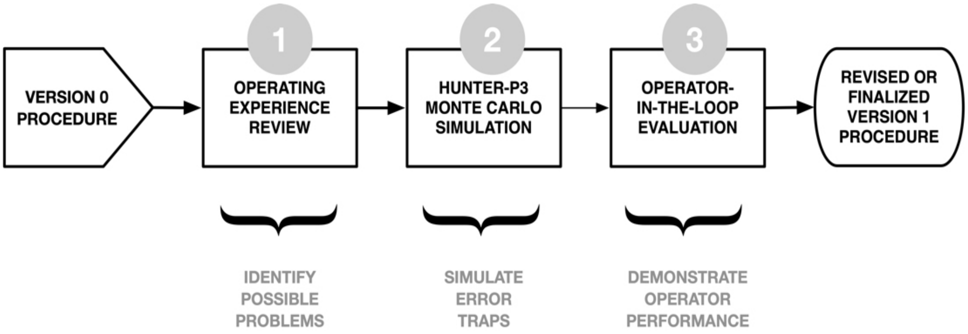

Figure 3: Three stages for validating version null procedures.

The goal of HUNTER-P3 is to identify overall error traps in the procedures and in the operator use of the procedures to allow procedure writers to refine the procedure as necessary prior to deployment. This approach provides a suitable procedure analysis tool between operating experience reviews and empirical evaluations. After operating experience reviews identify possible problem areas, these can be considered in the HUNTER-P3 modeling. If HUNTER-P3 reveals problem areas with procedures, this process serves as a screening tool to identify those use cases that should be explored further through operator studies. This process is illustrated in Figure 3. A graded approach in which not all three phases are performed is possible, and there may be circumstances when performing the HUNTER-P3 evaluation is sufficient to preclude operator studies because it sufficiently identifies procedural issues to be resolved.

As illustrated previously in Figure 1, HUNTER-P3 models are informed by human performance data. At first glance, it may seem paradoxical to suggest HUNTER-P3 modeling prior to operator-in-the-loop evaluation. To forego this step, assume that the HUNTER-P3 model is mature, meaning a phase has been completed already to calibrate individual plant operational characteristics to HUNTER and can be generalized to future scenarios. If separate human performance studies are necessary prior to running HUNTER, Stages 2 and 3 in Figure 3 may need to be reversed. However, once a HUNTER model is calibrated to a plant, it should not be necessary to have ongoing human performance studies for the purpose of model building.

A limitation of this approach is that most plant training simulators are not currently capable of running much faster than real time. The purpose of a training simulator is to allow human interactions that closely follow plant changes. Timing precision contradicts accelerated operations, and plant simulators closely follow actual plant response timings. The Rancor Microworld Simulator runs in a so-called headless version that is not linked to external timing constraints but can pass along timing durations. While Rancor may run a particular plant function hundreds of times faster than real time, it logs the time required of the plant evolution, enabling HUNTER to respond as if the actual time had passed and allowing faster-than-real-time synchronous coupling with HUNTER. This consideration matters because the use of Monte Carlo simulation runs to capture the range of human performance in HUNTER may prove a dauntingly slow process with a plant's actual training simulator. A 20-minute scenario may be run 500 times in Rancor faster than real time in under a minute. In contrast, the same 500 runs may take 10,000 minutes (ca. 167 hours) when performed with an unoptimized full-scope plant simulator. Parallelizing simulator installations for multi-core simultaneous execution or optimizing the simulator vendor's software code to run faster than real time can overcome this limitation.

DISCUSSION

While HUNTER was developed generally as a tool for dynamic HRA—meaning a tool to support risk analysis—the application of HUNTER-P3 to procedures illustrates the strong potential to benefit non-quantitative risk uses. Procedures are essential to the safe operation of nuclear power plants (and, indeed, to many other industries). As digital control systems are introduced into existing control rooms, this potentially changes the concept of operations at these plants and requires updates to procedures. These updated or new procedures can be considered with quantitative HRA methods, but many of the methods lack nuance to differentiate the consequences of changes in actions prescribed by those procedures. The human error probabilities predicted for procedures may not change, because the risk related to general actions and safety impacts do not necessarily change just because the human-system interface has evolved. On a task execution level, the procedures do change, and the reliability of operators to perform procedural tasks has the potential to change in ways that may not be fully reflected in the risk analysis. HUNTER-P3 offers a tool that can augment existing methods of evaluating novel procedures. By anticipating the types of error traps that can occur at both the procedure and operator levels, it offers a unique solution to vetting and optimizing procedures. The outputs of HUNTER-P3 include traditional error measures, but as a dynamic modeling tool, it can also capture issues with the flow of the procedure that may not rise to overt errors but can impact optimal operations.

DISCLAIMER

An extended abstract related to this work first appeared in the proceedings of the 2023 European Safety and Reliability (ESREL) conference. This work of authorship was prepared as an account of work sponsored by Idaho National Laboratory (under Contract DE-AC07-05ID14517), an agency of the U.S. Government. Neither the U.S. Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

REFERENCES

Bell, D. E., Raiffa, H., & Tversky, A. (1988). Decision Making: Descriptive, Normative, and Prescriptive Interactions. Cambridge University Press.

Boring, R., Lew, R., & Ulrich, T. (2017). Advanced nuclear interface modeling environment (ANIME): A tool for developing human-computer interfaces for experimental process control systems. Lecture Notes in Computer Science, 10293, 3–15.

Boring, R., Ulrich, T., Ahn, J., Heo, Y., & Park, J. (2022). Software Implementation and Demonstration of the Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER), INL/RPT-22-66564. Idaho Falls: Idaho National Laboratory.

Boring, R. L., Ulrich, T. A., Lew, R., Kovesdi, C., & Al Rashdan, A. (2019). A comparison of operator preference and performance for analog versus digital turbine control systems in control room modernization. Nuclear Technology, 205, 507–523.

Gertman, D., Blackman, H., Marble, J., Byers, J., & Smith, C. (2005). The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883. Washington, DC: U.S. Nuclear Regulatory Commission.

Hollnagel, E. (2015). Why is work-as-imagined different from work-as-done? In R. L. Wears, E. Hollnagel, & J. Braithwaite (Eds.), Resilient Health Care: The resilience of everyday clinical work (Vol. 2, pp. 249–264). Ashgate. Ashgate Studies in Resilience Engineering

International Atomic Energy Agency. (2018). Operating Experience Feedback for Nuclear Installations, SSG-50.

Lew, R., Boring, R., & Ulrich, T. (2018). Transitioning nuclear power plant main control room from paper based procedures to computer based procedures. Proceedings of the Human Factors and Ergonomics Society, 1605–1609.

Lew, R., Ulrich, T. A., & Boring, R. L. (2022). Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) Demonstration: Part 2, Model Runs of Operational Scenarios, INL/RPT-22-70076. Idaho Falls: Idaho National Laboratory.

Park, J., Yang, T., Boring, R. L., Ulrich, T. A., & Kim, J. (2023). Analysis of human performance differences between students and operators when using the Rancor microworld simulator. Annals of Nuclear Energy, 180, Article 109502.

Ulrich, T. A., Lew, R., Werner, S., & Boring, R. L. (2017). Rancor: A Gamified Microworld Nuclear Power Plant Simulation for Engineering Psychology Research and Process Control Applications. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61, 398–402.