Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 2780–2790
https://doi.org/10.54941/ahfe1008502

The Vigilance Decrement in the Wild: How to Study and Design for Long-Term Process Monitoring

Kelly Dickerson,[1] Courtney Thomas,[1] and Ronald Boring[2]

[1]Oklo Inc., Santa Clara, CA, 95050, USA
[2]Idaho National Laboratory, Idaho Falls, ID, 83415, USA

ABSTRACT

Several industries have made shifts towards highly automated or autonomous operations, with driving, aviation, and maritime operations serving as the most common examples. These concepts of operations shift the human from active operator to supervisory controller. Supervisory control requires consistent system monitoring to ensure success if human intervention becomes necessary. New build commercial nuclear power plants are being developed to leverage significant automation and autonomy, using sophisticated distributed control systems to manage and control the plant while human monitors continuously observe ongoing processes and trends. Process monitoring requires vigilance, the ability to remain alert and attentive while observing to detect a low likelihood event (e.g., abnormal plant conditions). Vigilance decrement is the degradation on monitoring performance over time and has been a known human performance issue since 1948 (Mackworth, 1948) with over 3,000 publications specifically on vigilance decrement. Despite the size of the literature, this research has translated into limited practical interface and job design guidance. This paper reviews and summarizes the primary theoretical frameworks posited as explanations and outlines a simulator-based method for studying and characterizing vigilance in a scientifically rigorous way that can be directly translated into actionable design insights. The study methodology proposes a multidimensional measurement approach, using physiological, survey-, and performance-based measures to capture the magnitude and timeline of the vigilance decrement over the course of a 2-hour observation interval.

Keywords: Vigilance, Human systems integration, Nuclear process control, Attention, Automation

INTRODUCTION

Automation and autonomy are increasingly an operational constant in maritime operations, industrial process control, and nuclear power. It is well established that adding automation does not eliminate the need for human involvement but rather changes the role the human plays in safe operations of the system (Dickerson et al., 2025; Joe et al., 2026). For example, a nuclear operator may make few control inputs during steady-state operation while still monitoring whether reactor power and heat-removal systems remain

Received September 17, 2026; Revised October 23, 2026; Accepted November 2, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

properly coordinated. If an automated action fails or produces an unexpected response, the operator must understand both the plant condition and the automation's prior actions before deciding whether to intervene manually. The shift from active controller to supervisor of a highly reliable system aims to reduce routine operator workload, but increased automation often produces the opposite effect of increasing workload, despite a reduction in the number of tasks.

This counterintuitive effect of increased automation increasing workload is potentially due to the demand of sustaining attention over a prolonged period. Maintaining attention over a prolonged period of time is defined as vigilance (Davies & Parasuraman, 1982). The vigilance decrement has been a known issue in human factors since the original Mackworth (1948) clock studies, aimed at understanding radar operator performance. Vigilance performance deteriorates as time on task increases, with previous studies demonstrating that when task demands are high, a decrement can emerge as early as the first five minutes of the watch period (Helton et al., 1999; Helton et al., 2007). Time on task is not the only factor influencing the presence and onset of the vigilance decrement. Task attributes, the number and type of breaks, presence of other concurrent tasks, target event frequency, and task complexity all impact the rate and severity of the vigilance decrement.

While there is substantial research investigating the underlying mechanisms of the vigilance decrement, this information has not led to directly actionable guidance for designing human system interfaces (HSIs), automation schemes, staffing concepts, or task and work practices that support operators who monitor complex processes for hours at a time. This is an important gap because as Warm, Parasuraman, and Matthews (2008) highlight, loss of situation awareness due to limitations in maintaining attention over time are part of the root cause in many aviation and driving accidents (see also Hemmerich et al., 2025) and was a direct contributor to the high profile Patriot fratricide incident (see Hawley, 2006). These examples highlight the importance of designing HSIs to support vigilance in supervisory control systems in high consequence environments. Research in other fields shows that this problem is not unique to nuclear operations. Sailors working rotating Navy watch schedules showed slower responses and more lapses (Shattuck & Matsangas, 2016), while automated-driving studies found that very low workload can leave drivers less prepared when they need to intervene (Young & Stanton, 2002).

OBJECTIVES

This paper has three objectives, aimed at developing a systematic approach to studying vigilance while balancing the need for rigorous evaluation methods and a high level of ecological validity and realism. Note this project is not a systematic literature review and is a small portion of a larger project.

1. Summarize the major theoretical frameworks for the vigilance decrement and their design implications for occupations where monitoring is a core job requirement.

2. Identify the methodological limitations that impede translation of vigilance research into real-world requirements and guidelines.
3. Outline potential approaches to simulator-based techniques for characterizing vigilance using converging performance, physiological, and subjective measures.

VIGILANCE DECREMENT CAUSES

Frameworks describing the underlying mechanisms of vigilance decrement tend to fall into two broad categories: Resource depletion theories and underload theories.

Resource Depletion Theories

Sustained monitoring requires substantial cognitive and information processing resources (Warm et al., 2008). For example, Warm, Dember, and Hancock (1996) found that National Aeronautics and Space Administration Task Load Index (NASA-TLX, Hart & Staveland, 1988) workload scores were highly and linearly associated with time on task, suggesting progressive resource depletion. As the task unfolds across the monitoring period, the resources required to successfully complete that task become depleted. For example, Caggiano and Parasuraman (2004) investigated a multiple resource theory (MRT) account of vigilance decrement. Under MRT there is evidence that when a visio-spatial working memory task is paired with a vigilance task there is a greater performance decline than when the vigilance task is paired with a verbal working memory task. This suggests there is some independence between different types of cognitive resources and that resource depletion is limited to the specific pool of resources in use. However, Helton and Russell (2013) found that vigilance performance was negatively impacted by any concurrent working memory task, suggesting a strong interconnection between sustained attentional demand and working memory, rather than narrowly specific independent pools of cognitive resources. Thus, while both accounts support a resource depletion explanation, the manifestation and underlying causes of depletion differ. Both depletion and the resources depleted point to specific design considerations for developing HSI-based decrement mitigation systems. Design considerations supporting amelioration of resource depletion would focus on ways to reduce unnecessary attentional demands and provide opportunities for resource recovery. Observable symptoms of cognitive resource depletion include reduced detection performance, increased response time, and increased user-perceived workload, potentially with observable physiological changes.

Underload, Arousal, and Mind Wandering Theories

While prolonged monitoring has been shown to induce symptoms of resource depletion, an alternative viewpoint is that monitoring highly reliable autonomous systems results in long periods of little meaningful operator activity. The consequence of limited meaningful system engagement is lower stimulation, where low event rates reduce physiological arousal and

cognitive engagement. The symptom of reduced arousal and disengagement is that attention "drifts" or shifts away from the monitored process. The characteristics of this attentional drift tend to fall into two broad categories: structured drift and unstructured drift.

Structured drift captures events that pull attention away from the monitoring task because they are distracting in some way. Smallwood and Schooler (2015) describe mind wandering as the process where attention shifts from the ongoing (e.g., goal directed) task to thoughts that are not linked to (either internally or externally in the environment). The "unfocused" state of the attentional network is well established (see Raichle et al., 2001); however, this theory is only one of several possible underlying mechanisms driving vigilance decrement. Interestingly, there is a resource depletion account of mind wandering (Thompson, Besner, & Smilek, 2015). Thompson et al. propose via the resource-control model that preventing mind wandering through attentional control (e.g., executive control processes) also consumes cognitive resources. From this perspective, when the observer experiences underload they have to allocate more cognitive resources to preventing off task thoughts than they would if their task load was closer to optimal.

If the vigilance decrement is driven by one or more underload factors (e.g., low arousal, distraction, or boredom), the HSI, task, and job design will need to consider how to improve arousal and task engagement without inducing artificial workload. That is, the breaks in monitoring should be meaningful and necessary operational tasks that are perceptually and cognitively distinct from the monitoring content. This type of intervention would ensure a balanced and sustainable approach to optimal task loading. Analysis of the design of the monitoring HSI also needs to consider the salience of the information to-be-monitored relative to potential distractors in task-irrelevant portions of the HSI or broader work environment.

A related set of theories that integrate under and overload frameworks describe the vigilance decrement as a question of opportunity cost and motivation (Kurzban et al., 2013). Motivational theories acknowledge that sustained vigilance has a cognitive resource cost, and this cost does not occur in a value-neutral environment. Operators continuously allocate effort according to expected value, task demands, and availability of alternative activities. Monitoring for rare events provides little immediate and concrete reinforcement for sustained effort. It's all the fatigue with none of the reward.

Can One Group of Theories Better Inform a Design Solution?

None of these theories are necessarily mutually exclusive. The observable behaviors of reduced detection accuracy and increased workload can arise from multiple and possibly different underlying states, further increasing the complexity of designing an HSI that effectively supports sustained monitoring performance. In all likelihood the vigilance decrement is a complex interaction between time on task, workload, arousal, fatigue, engagement, expectancy, motivation, and event characteristics (see Cummings et al., 2013). Theoretical ambiguity and the complexity of what appears to be a simple behavioral phenomenon suggests that to navigate the translation from research to

actionable guidelines, researchers and practitioners need to collect multiple types of measurements to develop a holistic understanding of supporting sustained attention, rather than treating a decline in performance as a single event, driven by a single mechanism.

THE LABORATORY-TO-OPERATIONS GAP

The conventional laboratory-based approach to understanding vigilance involves creating simple visual or auditory displays. Participants monitor the display for some period of time (typically between 30 to 60 minutes, but longer vigils are also reported, e.g., Cummings et al., 2013) to detect an infrequent target event. Laboratory studies of vigilance often include discrete (non-ambiguous) target and non-target events. These study design parameters are important experimental controls; however, real-world process monitoring in nuclear energy is fundamentally different from the monitoring protocols typical in laboratory vigilance studies, as shown in Table 1.

Table 1: Why is process monitoring different?

Laboratory-Based Vigilance	Real-World Monitoring
Simple signals	Multiple interacting process variables
Discrete targets	Ambiguous or evolving abnormal states
Often limited/not representative duration	Hours of continuous monitoring
Fixed HSI	Dynamic HSIs and multiple displays
Artificial event rate	Very low to zero target events
Detection is a terminating event	Detection initiates further action

With the rapid development of advanced small modular reactors that leverage a high level of autonomy and in some cases autonomous control, it is important and timely to systematically probe the gap between conventional vigilance studies and the real-world conditions of monitoring in supervisor control systems. Bailey and Scerbo (2007) found that, as the complexity of the monitoring task increased, performance decreased while trust in automation increased, regardless of automation reliability. While trust in automation is useful for workload management and efficiency, over trust, particularly under fatigue conditions, can lead to errors with protentional safety impacts.

Translating Research Into Design Questions

While studies of vigilance effectively report the onset of the decrement (as early as 5 minutes (see Helton, et al., 2000) to 15 minutes (see Teichner, 1974), there remains limited information available on key operational considerations. Specifically, it is unknown if the onset timeline and magnitude of the degradation in performance are operationally relevant (in general, long term, or short term). Or, research does not pinpoint if the operator is actively aware and compensating for resource depletion and/or fatigue and how the HSI or an HSI-based intervention could be designed to support the

operator's management of their monitoring performance. These fundamental questions can be further translated into direct design questions:

1. When does *meaningful* performance degradation begin?
2. How quickly does the performance quality degrade?
3. What aspects of monitoring degrade first?
4. Can physiological, subjective, or other measures detect degradation *before* a performance failure?
5. Which tasks or HSI characteristics accelerate or mitigate the vigilance decrement?
6. What interventions effectively restore performance?

DEVELOPING A DESIGN FOCUSED VIGILANCE STUDY

This section summarizes key details of the methods and rationale for a future vigilance study that prioritizes operational realism while balancing the need for a scientifically valid and statistically rigorous design. The operational context is prolonged monitoring of a highly automated distributed control system that manages reactor parameters and power production for a commercial sodium fast reactor.

Study Objectives

Occupations that require prolonged monitoring must consider the cognitive impact of that aspect of personnel job design as they develop and evaluate the function allocation, task analysis, and job design. Understanding the required cognitive capacity for the monitoring portion of crew tasking will uncover the available resources for all other control room tasks. These two pieces of information (capacity used by monitoring and capacity available for additional tasking), when combined with the functional requirements analysis and function allocation, inform definition of crew roles and the optimal crew size. This data driven approach, particularly when used iteratively, starting during the conceptual design phase, ensures that later in system development (such as during the integrated system validation) there are no staffing-related surprises due to under-estimation of the cognitive demands associated with monitoring. With those engineering goals and the specific operational context in mind, the proposed vigilance study must:

1. Characterize the magnitude of the vigilance decrement during a 2-hour watch period.
2. Determine the onset of a measurable degradation in performance.
3. Establish the rate of degradation over time.

The proposed study focuses exclusively on normal operations, since that is most representative of day-to-day monitoring. Participants will monitor the automated plant and respond to parameter deviations using the displays and procedures expected during routine operations. With a normal operational baseline established, a later study can expand the evaluation into abnormal and emergency operating conditions. Understanding these three parameters

(magnitude, onset, and rate of degradation) of the vigilance decrement under normal operational conditions will inform the determination of operational consequences, the urgency in developing a mitigation approach, and the characteristics of a successful design focused mitigation.

Simulated Monitoring

To meet the study objectives we must be able to closely mimic realistic dynamic monitoring conditions during normal power generation operations. Any deviations from the real world need to be evaluated for their potential to impact the observed performance. Additionally, the nuclear industry is highly regulated and the regulatory body, the U.S. Nuclear Regulatory Commission, requires plant referenced simulators used for validation testing, training, and licensing (see 10 CFR 55.46 for description of plant reference simulators), further establishing the importance of a high degree of operational realism.

Some vigilance tasks employ artificial secondary tasks that are designed to place additional strain on working memory (e.g., N-back, in which the participant compares the current stimulus to one *n*-times back) as a means of understanding cognitive capacity. While including additional performance measures is useful, it also changes the nature of the work the study is intended to observe. The inclusion of a secondary task and its impact on study realism must be carefully evaluated. Use of physiological- and survey-based measures may be a better alternative if task or situational realism are the highest priority. Additionally, many monitoring tasks require monitoring multiple simultaneous processes, suggesting realistically there may be no need to implement surrogate secondary tasks.

Participant Selection

It is important to evaluate if operators are needed or if realistic performance can be elicited from general population participants. In general it is preferred to use members of the operational community and provide a level of pre-testing training that is appropriate for the tasks they will perform. Using general population participants for immersive and highly realistic scenario-based studies can lead to poorer than expected performance because non-expert participants can become overloaded by the complexity of the information (Jones, 2021). For basic principles work aimed at developing or validating a theoretical framework, simplified studies and general population samples are the ideal standard to ensure generalizability and eliminate confounds. When conducting research for design purposes of specific systems like control room HSIs, using a general population is less of a focus due to the specialization of nuclear personnel. Nuclear personnel, for example, have extensive training on monitoring, and it is important to understand and calibrate performance to the exact types of vigilance decrement they experience. This study therefore aims to recruit 8 to 10 nuclear process control experts with experience working with a similar control system at a sister plant.

The exact number of participants selected depends on the final study design, number of trials, and the practical consideration of participant availability. Human factors teams should plan to conduct exploratory power analyses

early in the evaluation development process to understand their trade space between design and sample size. Studies with the nuclear operational community are often statistically underpowered because of the small size of the reactor operator population and their limited availability to support studies. To some extent smaller sample size can be offset by increasing the number of observations collected from each individual and reducing the complexity of the study design.

Scenarios and Target Events

Scenarios should be representative of normal operating conditions, rather than catastrophic accident scenarios. For example, the proposed study under development uses scenarios focused on: (1) full power steady state operations and (2) low power decay heat removal. Each of these scenarios are representative of the demands of day-to-day operations. Within the scenarios, key parameters can be placed just outside of an acceptable range and serve as target events. Awareness of these deviations can be evaluated both with and without auditory and/or visual alarms. Studies of vigilance using real world scenarios can also introduce design features aimed at improving situation awareness (e.g., parameter highlighting) or improving crew awareness of their performance (e.g., reminders to take breaks or physiological-based attentional drift notifications).

To further establish realism, the study team will conduct a crew brief. A crew brief is a short meeting at shift change or during important events where a shift supervisor will provide relevant details about the plant, controls, weather, planned testing, etc. The pre-scenario crew brief serves the purpose of informing the participants of the mission for the scenario (e.g., generate power or observe heat removal) and the critical parameters to monitor for the duration of the shift. Participants will be informed during the crew brief that one of the parameters is a particular concern and that it should be monitored closely. Including this information in the briefing establishes awareness of the target event in a naturalistic way.

One aspect of the traditional vigilance paradigm that is challenging to implement with a high degree of realism is ensuring a sufficient number of targets to fully evaluate the participant's performance. To detect a vigilance decrement rare targets must occur, but to achieve realistic performance, representative of normal operational conditions, targets cannot be high salience or otherwise extraordinary. This makes target selection both difficult and critically important. When considering target selection it is helpful to reflect on:

1. Which important operational or safety parameters are available.
2. Which parameters naturally vary and can they be placed out of range without creating a major alarm or incident.
3. Which established response procedures are available for investigating the out of range parameter, once detected.

MEASURING PERFORMANCE

One limitation of the laboratory studies that drive vigilance theories is that they tend to rely on performance and physiological measures and neglect to include common survey-based HFE measures, such as the NASA-TLX (Hart & Staveland, 1988). The proposed study uses three types of measurements in part to provide a greater volume and diversity of data from a small participant sample and to enable a triangulation or convergence analysis to increase confidence in the observation of a ground truth for performance. The measurement categories are (1) performance, (2) physiology (e.g., eye tracking [dwell time, time to first fixation, gaze direction, saccade velocity] and heart rate variability), and (3) surveys.

Performance and physiological measures will be analyzed both at the scenario level and as a function of time on task. Each scenario will be decomposed into 10 minute bins, and changes in performance and physiological response to the task will be trended across each 2-hour watch period. The threshold for acceptable vigilance will be coupled to overall performance measures. Performance decrement is the point at which vigilance dips and corresponds to increases in operational errors or drift from optimal plant performance.

Performance will be measured in terms of hits, misses, and false alarms. Response times will be measured based on time-to-detection and time-to-respond (if a response is required). A hit is correctly indicating that a target is present. A miss is failing to indicate a target is present while a target is present. A false alarm is indicating that a target is present when the target is absent.

Two physiological measures will be collected: eye tracking and heart rate. From these several workload and fatigue related measures will be extracted. Subjective workload will be evaluated using NASA-TLX. Since this study is the first in a series of human performance studies for an HSI in a new-build plant, the system usability survey (SUS, Brooke,. 1996; Brooke, 2013) will also be collected.

Individual difference data will also be collected to aid in interpretation of potential outliers. Individual difference variables include prior night's sleep duration and quality, the time of day of the assessment, caffeine consumption, age, and operational experience.

FROM STATISTICAL TO DESIGN SIGNIFICANCE

The objective of this study is not to prove that vigilance decrement is present during prolonged engagement in supervisory control tasks. We know that there will be some degradation in performance as time on task increases. The objective is to determine when performance becomes unacceptable from an operational perspective, the conditions that influence that level of degradation, the measurable indicators of unacceptable performance, and using that information to support discovery of effective design interventions to mitigate, delay, or eliminate the vigilance decrement.

ACKNOWLEDGMENT

The views expressed in this document represent those of the authors. Battelle Energy Alliance manages INL for the U.S. Department of Energy's Office of Nuclear Energy under contract number DE-AC07-05ID14517.

REFERENCES

Bailey, N.R. and Scerbo, M.W., 2007. Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. Theoretical Issues in Ergonomics Science, 8(4), pp. 321–348.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), Usability Evaluation in Industry. London: Taylor and Francis.

Brooke, J., 2013. SUS: a retrospective. Journal of usability studies, 8(2), p. 29.

Caggiano, D.M. and Parasuraman, R., 2004. The role of memory representation in the vigilance decrement. Psychonomic bulletin & review, 11(5), pp. 932–937.

Cummings, M.L., Mastracchio, C., Thornburg, K.M. and Mkrtchyan, A., 2013. Boredom and distraction in multiple unmanned vehicle supervisory control. Interacting with computers, 25(1), pp. 34–47.

Davies, D.R., Parasuraman. R.(1982). The psychology of vigilance.

Dember, W.N., 1996. Joel S. Warm William N. Dember University of Cincinnati Peter A. Hancock University of Minnesota. Automation and Human Performance: Theory and Applications, p. 183.

Dickerson, K., Watkins, H., Sparks, D., Hughes Green, N. and Morrow, S., 2025. What Can the History of Function Allocation Tell Us About the Role of Automation in New Nuclear Power Plants?. Energies, 19(1), p. 220.

Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in psychology (Vol. 52, pp. 139–183). North-holland.

Hawley, J.K., 2006. Patriot fratricides: The human dimension lessons of Operation Iraqi Freedom. FA JOURNAL, 11(1), p.18.

Helton, W.S., Dember, W.N., Warm, J.S. and Matthews, G., 1999. Optimism, pessimism, and false failure feedback: Effects on vigilance performance. Current Psychology, 18(4), pp. 311–325.

Helton, W.S., Hollander, T.D., Warm, J.S., Tripp, L.D., Parsons, K., Matthews, G., Dember, W.N., Parasuraman, R. and Hancock, P.A., 2007. The abbreviated vigilance task and cerebral hemodynamics. Journal of Clinical and Experimental Neuropsychology, 29(5), pp. 545–552.

Helton, W.S. and Russell, P.N., 2013. Visuospatial and verbal working memory load: effects on visuospatial vigilance. Experimental brain research, 224(3), pp. 429–436.

Hemmerich, K., Luna, F.G., Martín-Arévalo, E. and Lupiáñez, J., 2025. Understanding vigilance and its decrement: theoretical, contextual, and neural insights. Frontiers in Cognition, 4, p. 1617561.

Joe, J.C., Kovesdi, C., Mohon, J., Boring, R., Le Blanc, K., Morrow, S. and Hughes-Green, N. (2026) Function allocation: Human factors engineering technical basis and recommendations for guidance. RIL 2026–05. Washington, DC: U.S. Nuclear Regulatory Commission.

Jones, M., 2021. Effects of simulation fidelity on learning transfer. Journal of Educational Informatics, 2(1), pp. 24–34.

Kurzban, R., Duckworth, A., Kable, J.W. and Myers, J., 2013. An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36(6), pp. 661–679.

Mackworth, N.H., 1948. The breakdown of vigilance during prolonged visual search. Quarterly journal of experimental psychology, 1(1), pp. 6–21.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A. and Shulman, G.L., 2001. A default mode of brain function. *Proceedings of the national academy of sciences*, *98*(2), pp. 676–682.

Shattuck, N.L. and Matsangas, P., 2016. Operational assessment of the 5-h on/10-h off watchstanding schedule on a US Navy ship: Sleep patterns, mood and psychomotor vigilance performance of crewmembers in the nuclear reactor department. *Ergonomics*, 59(5), pp. 657–664.

Smallwood, J. and Schooler, J.W., 2015. The science of mind wandering: Empirically navigating the stream of consciousness. Annual review of psychology, 66(1), pp. 487–518.

Thomson, D.R., Besner, D. and Smilek, D., 2015. A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. Perspectives on psychological science, 10(1), pp. 82–96.

U.S. Nuclear Regulatory Commission (2019) 10 CFR § 55.46 - Simulation facilities. Washington, D.C.: Office of the Federal Register. Available at: GovInfo 10 CFR 55.46 Details [Accessed 25 August 2026]

Warm, J.S., Parasuraman, R. and Matthews, G., 2008. Vigilance requires hard mental work and is stressful. Human factors, 50(3), pp. 433–441.

Young, M.S. and Stanton, N.A., 2002. Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors*, 44(3), pp. 365–375.