INL/CON-24-78496-Revision-0

A Comparison of Three Types of Computer-Based Procedures: An Experiment Using the Rancor Microworld Simulator

November 2023

Changing the World's Energy Future

Anna Catherine Hall, Ronald Laurids Boring PhD, Thomas A Ulrich, Roger Lew, Michelle Velazquez, Jinding Xing, Tim Whiting, Georgios Michail Makrakis

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/CON-24-78496-Revision-0

A Comparison of Three Types of Computer-Based Procedures: An Experiment Using the Rancor Microworld Simulator

Anna Catherine Hall, Ronald Laurids Boring PhD, Thomas A Ulrich, Roger Lew, Michelle Velazquez, Jinding Xing, Tim Whiting, Georgios Michail Makrakis

November 2023

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2023, Vol. 67(1) 2552–2557

DOI: 10.1177/21695067231205507
journals.sagepub.com/home/pro
Sage

A Comparison of Three Types of Computer-Based Procedures: An Experiment Using the Rancor Microworld Simulator

Anna Hall[1], Ronald L. Boring[1], Thomas A. Ulrich[1], Roger Lew[2], Michelle Velazquez[3], Jinding Xing[4], Tim Whiting[5], and Georgios Michail Makrakis[6]

Abstract

The nuclear power industry has historically used paper-based procedures, but a shift towards computer-based procedures (CBPs) has the potential to reduce human errors, alleviate mental workload, and improve work performance. Twenty-seven participants were randomly assigned to one of three CBP types using the Rancor Microworld Simulator, and each performed two different types of operational scenarios (startup and loss of feedwater). The three CBP types varied in levels of digitalization. It was hypothesized that there would be less favorable impressions of the most basic procedures, and that these would demonstrate lower usability than types with greater digitalization. Overall, our predictions were partially supported with some interesting caveats, specifically with some performance benefits for CBPs that provided indicators but not embedded controls. We discuss our findings in terms of optimal levels of digitalization/automation within nuclear operations and suggest pathways for future directions.

Keywords

Human-system integration, Interface evaluation, Usability, System design and analysis

Introduction

The United States' fleet of nuclear power plants (NPPs) is of great importance to the nation's energy security, providing approximately 20% of all electricity. While existing plants have been rigorously maintained and operations are safe and functional, continued maintenance is expensive, and securing legacy replacement components is becoming difficult. Thus, many plants are making capital investments in upgraded instrumentation and control systems.

One upgrade involves a shift away from using paper-based procedures towards computer-based procedures (CBPs). The purpose of a procedure is to guide operators through a particular task – they are expected to follow it precisely and without deviation from the instructions. A 1995 report from the U.S. Nuclear Regulatory Commission (NRC) revealed that problems following paper-based procedures contributed to the majority of reportable events (NRC, 1995). The report highlighted several challenges with paper-based procedures, including identifying the correct procedure to follow, the stress involved with following multiple nested procedures in emergency scenarios, and issues with divided attention.

Modernizing procedures includes the integration of electronics devices and machines, such as computers, into an environment where tasks were originally executed by humans (Parasuraman & Riley, 1997). Using digital tools can help lower the number of human errors produced by alleviating some of the operator's mental workload, which can directly impact performance (Hwang et al., 2008). The Institute of Electrical and Electronics Engineers (IEEE) Standard 1786

[1]Human Factors and Reliability Department, Idaho National Laboratory, Idaho Falls, Idaho, USA
[2]Virtual Technology and Design Program, University of Idaho, Moscow, Idaho, USA
[3]Psychology Department Idaho State University Pocatello, Idaho, USA
[4]Civil and Environmental Engineering, Carnegie Mellon University Pittsburgh, Pennsylvania, USA
[5]Computer Science Department, Brigham Young University Provo, Utah, USA
[6]Computer Science Department, University of Idaho, Idaho Falls, Idaho, USA

Corresponding Author:
Anna Hall, Human Factors and Reliability Department Idaho National Laboratory, PO Box 1625, Idaho Falls, Idaho, 83415, USA.
Email: anna.hall@inl.gov

classifies CBPs into three types, depending on the amount of digitalization integrated into the procedure (IEEE, 2022). By the IEEE's definition, a Type I procedure closely resembles a traditional paper-based procedure in that it displays the instructions on a computer screen. Type II has more capabilities and can additionally display process data and step logic, visualize results, and provide access links to displays and soft controls that reside on a separate system. Type III has the additional capability to automatically carry out sequences in the procedure and has embedded soft control features. Unlike Types I and II, Type III procedures can manipulate the plant directly from within the procedure instructions.

While a great deal of research by human factors scientists has investigated CBPs in terms of hybrid control room needs (Le Blanc et al., 2015), use in field activities (Le Blanc et al., 2012; Oxstrand & Le Blanc, 2018), and frameworks for implementation (Lew et al., 2018), a gap remains in our knowledge as to how these different procedure types impact usability and user impressions. Thus, the current experiment examined these outcomes using the Rancor Microworld Simulator (Ulrich et al., 2017). It was hypothesized that participants would have less favorable impressions of the most basic Type I procedures, and that these would demonstrate lower usability than Types II or III.

Method

Participants

Twenty-seven individuals (19 males, 8 females) participated in the study (Mean age = 27.85, *SD* = 1.41; Range = 21 to 62 years), with 25 participants coming from the Idaho National Laboratory (INL) Intern Enrichment Program. Nineteen participants reported that they did not have previous experience with nuclear operations. The INL Institutional Review Board approved the research (INL000172) and signed consent was received from the participants.

Experimental Design

We used a between-subjects experimental design with one independent variable: CBP-type with three levels (Type I ($n = 10$), Type II ($n = 8$), and Type III ($n = 9$)). To increase ecological validity, as is typical with human factors studies in nuclear operations, participants each completed two different scenarios: normal operations (startup) and abnormal event (loss of feedwater). The startup scenario required participants to use a procedure that would start up the reactor (bring the reactor to criticality and begin load-following). To complete this scenario successfully, the reactor must not be tripped (shut down). The loss of feedwater scenario required participants to carry out a fault diagnosis and then shut down the reactor. To complete this scenario successfully, the reactor must safely be shut down within the time limit.

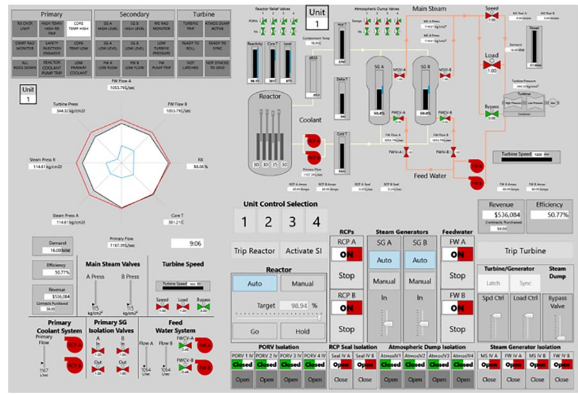

Figure 1. The Rancor Microworld Simulator, depicting the process overview (left), P&ID (upper right), and controls (lower right).

Rancor Microworld Simulator

Rancor is a simplified simulator that mimics both everyday operations and emergency situations performed in an NPP. It was developed at INL and the University of Idaho (Ulrich et al., 2017) by human factors scientists and contains essential NPP components and systems such as the reactor core, control rods, reactor cooling and feedwater pumps, bypass and load valves, steam generators, and turbine. The interface consists of three areas (see Figure 1). The Piping and Instrumentation Diagram (P&ID) depicts the relationship between piping, process equipment, instrumentation, and control devices. This displays the state of the plant. The Overview area collates much of the important information from the P&ID and puts it in one convenient location. This visual area also contains the alarm panel. The Control area allows the user to control the plant. These interface elements may be joined as a single screen or split across separate windows or screens.

Importantly, Rancor allows both operations experts and novices without operations experience to effectively monitor and control a simulated NPP to a proficient level. It monitors and records the status of all plant parameters as well as any actions taken by participants in each step of a procedure. Several studies have employed Rancor to successfully collect human factors and human reliability data involving nuclear operations. Recent studies have compared student and licensed operators and demonstrated good generalizability of findings from student participants (Park et al., 2022).

Measures

Rancor Simulator Logs. The simulator log data comprised the following:

1. Generated revenue (calculated as the difference in dollars between revenue at the beginning and end of each scenario).
2. Proportion of time in scenario with core temperature within range (between 400 and 650 degrees).
3. Actions performed (mean per scenario).
4. Scenario completion time (mean time in seconds).

Surveys. The National Aeronautics and Space Administration Task Load Index (NASA-TLX) measures subjective mental workload (Hart & Staveland, 1988). Participants were asked to rate how they felt completing each task on a 10-point scale. A higher score is indicative of a higher level of mental workload.

The Situation Awareness Rating Technique (3D-SART; Taylor, 1990) assesses attentional resources demand, attentional resources supply, and situation understanding on a 7-point Likert scale. The following formula is used to calculate the overall situation awareness (SA) score: SA = U – (D – S), whereby U represents *understanding*, D represents *demand*, and S represents *supply*. Higher scores are indicative of a higher level of situation awareness.

Using a 10-point Likert scale, participants rated the extent to which the procedures helped. Using a forced choice "yes / no" response, participants indicated whether they had been successful in completing the procedure and whether they had tripped the plant. Finally, participants indicated how much they liked the CBPs (10-point Likert scale).

Experimental Procedure

Participants were randomly assigned to one CBP-type (Type I, Type II, or Type III). They provided informed consent, completed a brief demographic survey, and watched a Rancor Microworld Simulator training video specific to their CBP-type group assignment. The video explained how nuclear power is generated, introduced the Rancor interface, the purpose of using procedures in nuclear operations, and the study's goals. The video then went through the startup scenario using the procedure in Rancor. Participants had three minutes to practice. The same occurred for the loss of feedwater scenario. This was followed by three trials of each scenario, after which the NASA-TLX, the 3D-SART, and other post-scenario items were completed. At the end of the study, participants indicated how much they liked the CBPs. Finally, they were debriefed and thanked for their time.

Data Analysis

The frequency data were analyzed using Chi-square goodness-of-fit tests. All other data were subjected to one-way Analyses of Variances (ANOVAs) to compare the effects of CBP-type on usability and user impressions. When appropriate, non-event and event scenarios were analyzed separately.

Results

Table 1 provides the results of the one-way ANOVAs. All post-hoc comparisons were conducted using Tukey- Kramer tests. Given the nature of different scenario types on the Rancor simulator log data, these were analyzed separately. For the startup scenarios, there was a significant main effect of CBP-type on Generated revenue: Type II CBP generated more revenue than Type I ($p = .022$). For the loss of feedwater scenarios, there was a significant main effect of CBP-type on actions performed: Type I CBP resulted in more actions performed than Type III ($p = .032$). There were no other significant findings for any other usability metric with the Rancor software.

For the survey data, there was a significant main effect of CBP-type on the NASA-TLX; Type I produced higher mental workload than Types II ($p = .038$) and III ($p < .001$). These results were similar for both scenarios. Overall, there was a non-significant trend for 3D-SART, but for the startup scenarios only, Type II CBP generated greater situation awareness than Type III ($p = .035$). There was no difference across CBP-types for liking ratings. However, Type I was rated less helpful than the others; this was driven by the startup scenario in which Type I produced lower ratings than Type II ($p = .018$) and Type III ($p = .016$).

Type I CBPs produced a greater likelihood of failure to complete the scenarios compared to Types II and III (χ^2 (2,150) = 8.57, $p = .014$) (Figure 2), driven by the startup (χ^2 (2,80) = 12.04, $p = .002$) but not loss of feedwater scenarios. There was also less success completing the startup scenarios overall (χ^2 (1,150) = 16.38, $p < .001$), likely a result of the lengthy duration involved with this procedure. Overall, the plant was tripped less with Type II CBPs compared to Type I and Type III, (χ^2 (2,148) = 6.41, $p = .040$) (see Figure 3). Finally, there was no significant difference in plant trips between startup and loss of feedwater scenarios.

Discussion

This is the first experiment to test usability and user impressions of the three types of control room CBPs defined by IEEE-1786. There are important links to these procedure types and levels of automation. While Type I had no automation embedded within the procedure, Type II featured information automation, and Type III featured a limited degree of control automation. We hypothesized that Type I CBPs would demonstrate lower usability and elicit less favorable impressions compared to Types II or III. Our predictions were partially supported but with some interesting caveats. The Rancor simulator logs revealed that greater procedure digitalization produced better participant performance (more

Table 1. Survey and Rancor Microworld results as a function of CBP-Type.

	Computer-Based Procedure (CBP)									
	Type I			Type II			Type III			
	n	Mean	SE	n	Mean	SE	n	Mean	SE	*F*
Startup										
Generated revenue	30	35,072	21,867	24	124,141	24,448	27	85,608	23,050	*3.76**
% time with core T in range	30	0.38	0.04	24	0.52	0.05	27	0.39	0.04	*2.98*
Actions performed	30	39.7	4.8	24	54.7	5.4	27	52.4	5.1	*2.58*
Time to complete scenario	30	545.2	14.0	24	543.0	15.7	27	516.2	14.8	*1.21*
Loss of feedwater										
Generated revenue	22	104,084	6,743	20	85,762	7,072	22	82,602	6,743	*2.93*
% time with core T in range	22	0.14	0.02	20	0.13	0.02	22	0.12	0.02	*0.37*
Actions performed	22	54.3	5.6	20	39.4	5.87	22	33.9	5.6	*3.56**
Time to complete scenario	22	540.2	22.9	20	535.4	24.1	22	493.3	22.9	*1.25*
Both scenarios										
NASA-TLX	54	5.78	0.19	43	5.08	0.21	53	4.67	0.19	*9.00****
3D SART	54	5.57	0.24	44	6.14	0.27	53	5.23	0.25	*2.75*
Helpfulness of CBP	54	8.22	0.21	48	8.88	0.23	54	8.89	0.21	*3.10**
Liking of CBP	54	7.89	0.26	48	8.25	0.28	54	8.44	0.26	*1.15*

**p*<.05; ***p*<.01; ****p*<.001.

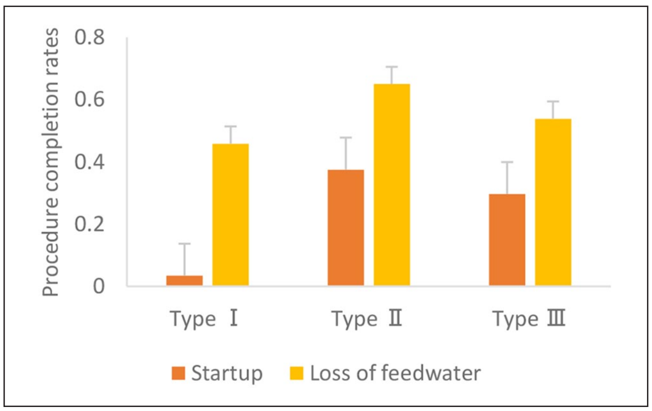

Figure 2. Proportion of trials finished.

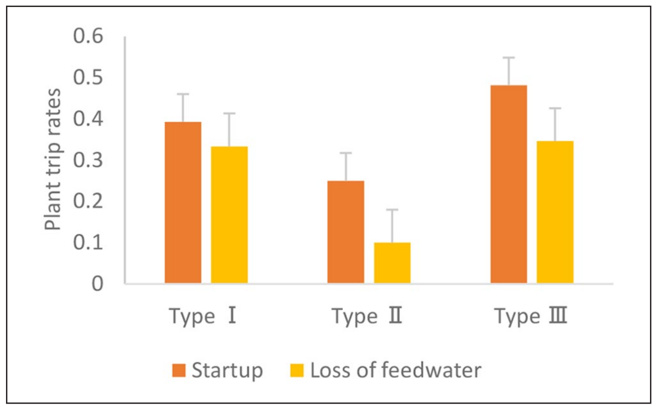

Figure 3. Proportion of trials with plant trips.

revenue and less actions taken), but that this depended on the type of scenario. Otherwise, there was little difference across CBP types. The survey data was more discerning in that there were clearer preferences for greater digitalization across both types of scenarios, but especially for startup.

Although the participants did not indicate any significant preference for one type of CBP over another, they did consider greater digitalization more helpful. Further, for the startup scenarios, Type II produced higher situation awareness compared to Type III. This is interesting considering that Type II also elicited the least amount of plant trips. It could be the case then, that the convenience of a highly digitalized procedure does not represent optimal conditions for users and instead causes a drop in awareness which affects performance, at least for the startup scenario. Additionally, Type II produced the highest completion rates. Taken together, these data suggest that Type II procedures, absent embedded soft controls as with Type III, but with the addition of plant process indicators and step logic compared to Type I, may represent the 'sweet spot' of computerized procedures, at least for control room operations.

This finding, an inverted U-shaped function, was surprising given the age range of participants who represent 'digital natives.' This term describes the generation that grew up in the last 25 years, with digital technology being a part of their entire lifetimes, surrounded by the internet and personal devices (Gani, 2018). Thus, it was expected that this younger cohort would prefer greater technology and more sophisticated levels of digitalization as with Type III CBPs. This expectation was borne out for mental workload results, which was reported to be lowest for Type III. Yet generally, the findings speak more to a manual / digital balance being favored. It may be the case that when perceived stakes are high, such as nuclear operations, individuals wish to retain some level of autonomy over control functions. This is in keeping with prior research that suggests existing NPP operators may feel discomfort relinquishing controls to technology, and a step-wise approach to modernized upgrades will likely curry more favor in the near term (Lew et al., 2017).

Nonetheless, the overall finding of the inverted U-shaped function mimics several other psychological processes in which performance or optimal conditions are represented in the middle. For example, it is well documented that exposure to some amounts of stress facilitates performance in domains such as organization (Anderson, 1976), work productivity (Wilke et al., 1985), and cognitive function (Sandi, 2013). This has also been demonstrated in animal research (Salehi et al., 2010). Too much or too little stress intensity results in a performance decline. This psychological principle may apply to the human factors research outlined here such that the digitalization level introduced to NPP control rooms must be carefully gauged somewhere in the middle for the greatest benefit. Certainly, this is in keeping with several prior works that highlight this necessary balance (Boring et al., 2019; Marras, 2010). Additionally, future research should consider how developmental aging impacts digitized systems and automation in CBPs, given different age-sensitive motivations to adopting new technologies (Morris & Venkatesh, 2000), and cognitive skillsets that may favor older operators in certain advanced automated protocols (Hall et al., 2022).

Strengths of this research include being the first study to experimentally test usability and impressions of the three different types of CBPs as defined by the IEEE. Given that CBPs present a pathway to NPP cost-savings, and by extension shoring up U.S. national energy security, these findings are critical to the industry. Another strength is the experimental design, allowing causal inferences to be made. Participants were randomly assigned to each of the three procedure types, balancing out any individual differences between group members. In addition, the experimental protocol required the participants to complete the scenarios multiple times, providing greater reliability in our findings.

However, these results must be considered within some limitations. First, our sample size was small and so the findings reported here should be considered preliminary. Second, the participants were novices, and not NPP operators, and we used a simplified version of a nuclear operations simulator which may have impacted the results. Future research should aim to have larger sample sizes and test operators and where possible, within a full scope NPP control room simulator. It would also be interesting to see if there are differences in the type of CBP preferred by novices and NPP operators. Furthermore, this would amplify our understanding of the feasibility and generalizability of CBPs, and whether our findings can be replicated under more realistic control room conditions.

Acknowledgments

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.

ORCID iDs

Tim Whiting https://orcid.org/0000-0003-4016-1071
Georgios Michail Makrakis https://orcid.org/0000-0002-1280-6568

References

Anderson, C. R. (1976). Coping behaviors as intervening mechanisms in the inverted-U stress-performance relationship. *Journal of Applied Psychology*, *61*(1), 30-34.

Boring, R. L., Ulrich, T. A., & Mortenson, T. J. (2019). Level-of-automation considerations for advanced reactor control rooms. *Proc. 11th Nuclear Plant Instrumentation, Control, and Human- Machine Interface Technologies*, 9-14.

Gani, S. A. (2016). Parenting digital natives: Cognitive, emotional, and social developmental challenges. In *International Conference on Education (ICE2) 2018: Education and Innovation in Science in the Digital Era* (pp. 870-880).

Hall, A., Boring, R. L., & Miyake, T. M. (2022). Cognitive aging as a human factor: Effects of age on human performance. *Nuclear Technology*, 1-15.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload*, 139–183.

Hwang, S. L., Yau, Y. J., Lin, Y. T., Chen, J. H., Huang, T. H., Yenn, T. C., & Hsu, C. C. (2008). Predicting work performance in nuclear power plants. *Safety science*, *46*(7), 1115-1124.

IEEE. (2022). *IEEE Guide for Human Factors Applications of Computerized Operating Procedure Systems (COPS) at Nuclear Power Generating Stations and Other Nuclear Facilities, IEEE- 1786-2022*. New York: IEEE.

Le Blanc, K. L., Oxstrand, J. H., & Joe, J. C. (2015). *Requirements for control room computer-based procedures for use in hybrid control rooms* (No. INL/EXT-15-35284). Idaho National Lab. (INL), Idaho Falls, ID (United States).

Le Blanc, K., Oxstrand, J. H., & Waicosky, T. (2012). Requirements for computer-based procedures for nuclear power plant field operators. Results from a qualitative study (No. IAEA-CN--194).

Lew, R., Boring, R. L., & Ulrich, T. A. (2018). Transitioning nuclear power plant main control room from paper-based procedures to computer based procedures. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1605–1609. https://doi.org/10.1177/1541931218621363

Lew, R., Ulrich, T. A., & Boring, R. L. (2021) 'Rancor Hybrid Energy System Microworld'. SAGE Publications CA: Los Angeles, CA, 64(1), pp. 1760–1764. doi: 10.1177/1071181320641426.

Marras, W. S. (2010). Bored beyond belief: How automation is boring us to distraction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(11), 758–758. https://doi.org/10.1177/154193121005401102

Morris, M.G., & Venkatesh, V. (2000). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel Psychology, 53*, *2*, 375-403.

NRC (1995). Evaluation of the computerized procedures manual II (COPMAI I). Oxstrand, J. H., & Le Blanc, K. (2012). Computer-based procedures for field workers in nuclear power plants: Development of a model of procedure usage and identification of requirements. *Idaho National Laboratory External Report*.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253.

Park, J., Boring, R.L., Ulrich, T.A., Lew, R., Lee, S., Park, B., & Kim, J. (2022). A framework to collect human reliability analysis data for nuclear power plants using a simplified simulator and student operators, *Reliability Engineering & System Safety*, *221*, Article 108326.

Salehi, B., Cordero, M. I., & Sandi, C. (2010). Learning under stress: the inverted-U-shape function revisited. *Learning & Memory*, *17*(10), 522-530.

Sandi, C. (2013). Stress and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 245-261.

Taylor, M. (1990). Situation awareness rating technique (SART): The development of a tool for aircrew systems design. *Situation Awareness in Aerospace Operations*, 3-17. Advisory Group for Aerospace Research & Development.

Tubbs-Cooley, H. L., Mara, C. A., Carle, A. C., & Gurses, A. P. (2018). The NASA Task Load Index as a measure of overall workload among neonatal, pediatric and adult intensive care nurses. *Intensive & Critical Care Nursing*, *46*, 64–69. https://doi.org/10.1016/j.iccn.2018.01.004

Ulrich, T.A., Lew, R.T., 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.

Wilke, P. K., Gmelch, W. H., & Lovrich Jr, N. P. (1985). Stress and productivity: Evidence of the inverted U function. *Public Productivity Review*, 342-356.