Exploring measures of workload, situation awareness, and task performance in the Main Control Room

Alexandra Fernandes*, Per Øivind Braarud

OECD Halden Reactor Project, Institute for Energy Technology, Os Alle 5, Halden 1777, Norway

Abstract

Performance based evaluations are commonly completed as part of the human factors verification and validation of control rooms of new builds or substantial modernizations in Nuclear Power Plants. The validation of the Main Control Room system in the nuclear industry entails the assessment of specific performance components such as plant performance, situation awareness, workload, task performance, and anthropometric and physiological factors. This evaluation is expected to take place in realistic contexts, simulating the plant’s operation and covering also abnormal and emergency operation situations. There is a vast set of measures to address these human performance components, and the relations among the different concepts and different measures are not clearly defined. The measures used to assess human performance can also vary significantly regarding the type of rater (self-rated, observer-rated, objective ratings), the type of scale (rating scales, correct/incorrect decisions, open answer, acceptability judgments), the moment of rating (online rating, end of scenario, scenario freezes), among others. In this paper we try to analyze these relationships between task performance, workload, and situation awareness. We explored the results of both self-rated, observer-rated, and objective measures. We explored the relations between and within measures, as well as the trends across the different assessments, affected by control room modifications. The results indicate a similar pattern in workload and situation awareness measures through evaluations; that the relations within self-rated measures appear to be clearer than between self-rated and observer-rated measures; and that there seems to be a trend for improved task performance results in the control room throughout the years.

* Corresponding author. Tel.: +47 465 43 479 .
E-mail address: alexandra.fernandes@hp.no
1. Assessing human performance

Human performance is a crucial aspect in control room evaluations in the nuclear industry. The disciplines of Human Factors, Cognitive Engineering and Ergonomics rely on the development, testing and enhancement of performance measures to fulfill their role in understanding, promoting, and explaining human behavior in complex systems. There are many factors that might affect, restrict, or enhance human performance and as such this is a multifaceted concept, comprising various dimensions from contextual and environmental influence, to high order internal cognitive processes of reasoning, as well as social and teamwork factors, basic perceptive and motor phenomena, training effects, amongst many other relevant dimensions.

Specifically in the nuclear industry, in order to license operation, regulations for main control room validation and verification require that human performance is assessed in the context of overall system performance – involving personnel, hardware and software systems and their integration [1]. Both physiological factors (such as vision or motor control) and cognitive factors (such as attention and memory) have an impact on the tasks that the personnel need to complete in the control room [2]. The physiological factors can be regarded as more easily observable and testable, being related with ergonomics and the physical capacity of the operators – instruments in reach, visibility from the usual sitting place, capacity to hear other team members from the working station, etc. Not being directly observable, the cognitive factors raise other challenges – what aspects of cognition are relevant for the tasks in the control room? Which ones are the most relevant? Do they contribute to the understanding of performance? Could they support predictions of future performance? These are central aspects in human performance and practitioners, researchers, and regulators have been making a significant effort to define and operationalize them [2], [3], [4]. We will give particular emphasis to the concepts of situation awareness, workload, and task performance, since they are particularly frequent in human factors, not only in the nuclear industry [5], but also in the military [6], aviation [7], railway [8], oil & gas [9], driving [10], or process industry [11].

Situation awareness refers to the individual/team perception of the surrounding environment and its elements at a certain point in time, including also an understanding of the current status and a prediction of a future situation or identification of a trend in the system [7]. This definition is generic and is applicable to different industries and tasks. The operationalization of the concept and the development of measures to evaluate it is however dependent on the specific contexts [12], with different elements of situation awareness being more emphasized in one domain than in others.

Workload, more specifically cognitive workload, refers to the amount of work that an individual can perform at a certain time, or more specifically, the amount information an individual can process or the mental effort required to complete a task [13]. Workload is commonly self-evaluated, reflecting an individual perception of a multidimensional concept, covering aspects related with mental and physical effort, frustration, or time constraints that go beyond the objective task difficulty and demands [14].

The overall idea of human performance in the control room relates to the accomplishment of a task according to predefined parameters of accuracy and efficiency. As such, the observable aspects of human performance are determined by the assessment of task performance, which can be objectively assessed by the completion of specific observable tasks. In NUREG-0711 [1], performance measures are presented in a hierarchal fashion, building up from physiological and anthropometric aspects to workload, then situation awareness, personnel task performance, and finally overall plant performance. As such, workload is implicitly presented as an influence factor in situation awareness, and situation awareness is believed to have a determinant role on personnel task performance. These are very relevant assumptions in verification and validation of control rooms in the nuclear industry and we will attempt to explore these relationships between concepts in this paper.

In the next sections we will layout the overall goals for the analysis presented in this paper, focusing on two central questions regarding the measures of workload and situation awareness used in control room evaluation: 1) What are they for? and 2) Are they useful?
1.1. What are the concepts of workload and situation awareness for?

As mentioned before, the relevance to assess constructs such as cognitive workload and situation awareness relies on their assumed impact on task performance. Understanding and predicting human performance is the ultimate goal for human factors analysis, and concepts about how people perceive, interact and understand their surrounding are important for their actions. As such, aspects such as workload and situation awareness are relevant to the extent that they provide insight into the operators’ internal model of how things work, and help to explain the subsequent decisions and actions [4]. The utility of assessing these aspects has relied essentially on the possible recommendations and adjustments that can be made to the working environment to facilitate the desired behavior and to prevent inadequate responses, for instance by improving the interface design, adding elements to a team, repositioning screens so that everyone can see them, amongst others.

1.2. Are the concepts of workload and situation awareness useful?

The concepts of situation awareness and workload were driven from a need to explain human performance not only through observable behavior, but also relying in cognitive processing aspects that are not directly observable. Nonetheless, the measurement, definition, and implication of assuming that these concepts have a direct effect on task performance can still be controversial [15].

A particularly relevant question rises from the correlations or connections between the measures of the different concepts [16]: can we assess situation awareness and workload separately? The definition of the two concepts seems to pinpoint clearly different aspects, however their impact on task performance might be combined, for instance assuming that situation awareness can have an effect on the time it takes to complete a goal, and that workload might contribute to pace the decision-making and search time available for situation awareness processes. Berggren et al. [6] have explored these connections through structural equation modeling and reported some preliminary results where a model that presents situation awareness as a mediator between the joint effects of workload and teamwork on performance seemed to be promising.

Overall, the collective effort to improve the models, definitions, and measurement of the cognitive dimensions related to task performance, namely workload and situation awareness, has resulted in a significant increase of the knowledge and understanding of human performance, and contributed to practical improvements for operators [4].

2. Method

In the nuclear industry Verification and Validation activities assess whether the control room follows human factors engineering design principles that allow the operators to efficiently complete their tasks, supporting the achievement of plant safety and operational goals [1]. A central part of Verification and Validation refers to the application of performance-based tests in full-scale simulators, enabling the testing in a “close-to-real” context that will allow the generalization of results to real-life performance in the plant (Integrated System Validation - ISV).

We will focus on a selected analysis of the measures/data collected in three ISV evaluations in a specific plant through a period of approximately 6 years. Evaluation 1 (E1) refers to the initial control room simulator, before any control room modernization; Evaluation 2 (E2) refers to data from the ISV performed right after a control room modernization in the new simulator (3 years later than E1); and Evaluation 3 (E3) reports data from an ISV evaluation done 3 years after E2 (and as such after some familiarization with the modernized simulator/control room), as a long-term validation plan to the new control room.

Having this long-term approach, allowed us to focus on the measurement correlations, and provided a bigger data sample using the same measures in different moments. The data we present here corresponds to the simulator evaluation of four scenarios (2 accident and 2 non-accident scenarios), repeated in three different moments of evaluation with different crews from the plant. According to the evaluation teams, these scenarios were considered very good to assess control room overall performance, since they were complex and demanding, allowing a realistic test of the set-up. The presented analysis corresponds to overall data, collected in all scenarios.
2.1.1. Workload measures

Workload was evaluated with the NASA Task Load IndeX – NASA-TLX (Hart & Staveland, 1988). This scale is a self-rated scale (subjective experience, rated by operators) and consists of six sub-scales – mental demand, physical demand, temporal demand, performance, effort, and frustration – evaluated in a rating scale with 11 points. The performance subscale is evaluated from “failure” (0) to “perfect” (100) and all the others are labeled from “low” (0) to “high” (100). A generic workload rating was also obtained from two external observers in a seven points rating scale.

2.1.2. Situation awareness measures

The Situation Awareness Control Room Inventory – SACRI [17] is a measure that allows the direct assessment of the operators’ knowledge of current, past, and future status of specific parameters in the plant (comparing answers to real status). The measure has 18 items and, for each one, the operators must answer whether a specific parameter has increased, decreased, or remained stable in the last period of time. A three-item questionnaire on situation awareness (covering information gathering, process status, and predictability of the process) was also presented to the operators to assess the subjective perception of obtained overview. Observers also rated in a 7 points scale the process status understanding of the operators and prediction of process development.

2.1.3. Task performance measures

The Operator Performance Assessment System – OPAS [18] was used in this evaluation providing a task-goal oriented performance score. This measure entails that a process expert analyses the test scenarios and the required activities in each one defining an ideal performance situation that serves as a comparison to the real performance of the crews during the test (proportion of completed tasks). This measure involves activities that relate both with information gathering and with direct actions on plant process. This is considered a direct measure since it entail the comparison of actual and ideal performance. Observers rated in a scale (rating scale with 7 points) teams’ task performance. An operator-rated performance measure was also presented, covering performance in crew’s tasks.

3. Results

The collected data had different sources and each measurement varied in the type of scale used for the rating. As such, to facilitate the analysis and interpretation, we chose to normalize all the data before the correlation analysis. The tables below show both the original and the normalized data used in this analysis.

3.1. Workload

The ANOVA analysis revealed no significant differences in the workload measurements in each of the evaluation moments: \( F(4, 184) = 1.45, p = .22 \). Figure 1 illustrates the results.
The data analysis revealed that there is a significant correlation between the observer and operator rating for workload ($\tau = .52, p < .05$). This correlation was observable in all moments of evaluation: E1 = .44, $p < .05$; E2 = .61, $p < .05$; and E3 = .50, $p < .05$.

Table 1. Workload values for each moment of evaluation (standard deviation in brackets).

| Observer Rating | Operator Rating (NASA - TLX) |
|-----------------|-------------------------------|
|                 | Original Score | Normalized Score | Original Score | Normalized Score |
| E1              | 5.45 (1.30)    | 0.61 (0.32)      | 37.97 (12.40)  | 0.54 (0.26)      |
| E2              | 5.34 (1.39)    | 0.59 (0.35)      | 38.00 (10.85)  | 0.54 (0.23)      |
| E3              | 5.45 (1.20)    | 0.61 (0.30)      | 34.11 (10.98)  | 0.46 (0.23)      |

3.2. Situation awareness

The analysis if the situation awareness evaluations in the three moments of evaluation revealed equivalent ratings in the measure for all measures in all moments: $F (6, 182) = 2.10, p = .06$. There is a trend for a lower rating in the evaluation done immediately after the modernization (E2) for operator-rated and objective measures, possibly showcasing less familiarity with the control room.

The ratings obtained in situation awareness assessment from the observers, operators, and from the objective evaluation did not seem to correlate with each other generally ($p = \text{n.s.}$), and considering each moment of evaluation only E2 revealed a significant correlation between observer and operator-rated situation awareness ($\tau = .27, p < .05$).

Table 2. Situation Awareness values for each moment of evaluation (standard deviation in brackets).

| Observer Rating | Operator Rating | Objective Evaluation (SACRI) |
|-----------------|-----------------|-------------------------------|
|                 | Original Score  | Normalized Score | Original Score | Normalized Score | Original Score  | Normalized Score |
| E1              | 4.97 (1.36)     | 0.59 (0.27)      | 5.27 (0.72)    | 0.59 (0.20)      | 0.79 (0.11)     | 0.75 (0.16)     |
| E2              | 4.89 (1.23)     | 0.58 (0.25)      | 4.85 (0.73)    | 0.47 (0.20)      | 0.71 (0.12)     | 0.64 (0.18)     |
| E3              | 4.94 (1.02)     | 0.59 (0.20)      | 5.20 (0.71)    | 0.57 (0.19)      | 0.73 (0.10)     | 0.66 (0.14)     |

Fig. 2. Situation Awareness measures in each evaluation moment.
3.3. Performance

The performance ratings showed a significant statistical difference across moments of evaluation \( [F (6, 182) = 2.22, p < .05] \). This contrast resulted from a significant difference between E2 and E3 in OPAS \( (t (32) = -2.24, p < .05) \) and the operator-rated performance \( (t (32) = -3.83, p < .01) \), favoring performance in E3. Although non-significant, there is an observable trend to better performance results in E3 than in the original control room (E1).

In the performance measures it is possible to observe a correlation between the objective measure and both the observer ratings \( (\tau = .35, p < .05) \) and the operator ratings \( (\tau = .15, p < .05) \). However, there is no correlation between the ratings of the operators and the observers. In the evaluation moment analysis a correlation is observed only for observer-rated and objective measures in moment E1 \( (\tau = .54, p < .05) \) and E2 \( (\tau = .25, p < .05) \).

| Observer Rating | Operator Rating | Objective Evaluation (DOS – OPAS) |
|-----------------|-----------------|----------------------------------|
|                 | Original Score  | Normalized Score                 | Original Score | Normalized Score |
| E1              | 4.83 (1.48)     | 0.61 (0.27)                      | 5.84 (0.80)    | 0.65 (0.24)      |
|                 | 81.51 (12.34)   | 0.66 (0.23)                      |                |                  |
| E2              | 4.73 (1.20)     | 0.59 (0.22)                      | 5.56 (0.68)    | 0.57 (0.20)      |
|                 | 76.67 (12.55)   | 0.57 (0.23)                      |                |                  |
| E3              | 5.06 (1.97)     | 0.65 (0.19)                      | 6.09 (0.39)    | 0.73 (0.12)      |
|                 | 82.92 (9.67)    | 0.68 (0.18)                      |                |                  |

3.4. Relations between measures

3.4.1. Relation between observer ratings

The correlations amongst the observer ratings revealed a connection between the rating of situation awareness and performance \( (\tau = .69, p < .05) \) and situation awareness and workload \( (\tau = -.14, p < .05) \). The correlations between workload and performance were non-significant.

3.4.2. Relation between operator ratings

The operator ratings between situation awareness and workload were significantly connected \( (\tau = -.34, p < .05) \), with lower workload values being associated with higher situation awareness ratings; as well as the ratings of situation awareness and performance \( (\tau = .48, p < .05) \) with a positive correlation. The Kendall’s tau correlation was also significant for the ratings of workload and performance \( (\tau = -.31, p < .05) \), with better performance ratings linked to lower workload values.
3.4.3. Predicting performance measures

We intended to analyze the predictive value of the obtained situation awareness and workload ratings in the prediction of performance. The results from the multiple regression of observer, operator and objective ratings of workload and situation awareness on the performance results are presented in Table 4.

For the observer-rated performance it was possible to find that about 71% of its variance is explained by the workload and situation awareness measures \( F (5, 90) = 43.27, p < .01, \text{error} = .13 \). The observer-rated situation awareness was the best single predictor for observer-rated performance, with \( \beta = .84 \) \( t (90) = 13.93, p < .01 \).

The multiple regressions showed that for the operator-rated performance, around 38% of variance could be explained by the independent variables \( F (5, 90) = 10.98, p < .01, \text{error} = .16 \). In this case, the best individual predictor was the operator-rated situation awareness, with \( \beta = .48 \) \( t (90) = 4.80, p < .01 \).

Finally, for the objective measure of performance OPAS, about 35% of the variance was explained \( F (5, 90) = 9.74, p < .01, \text{error} = .18 \), with the observer-rated situation awareness being the main predictor with \( \beta = .49 \) \( t (90) = 5.50, p < .01 \).

All the other independent variables were not able to show a relevant predictive value of the variance in the three target performance measures.

| Observer-rated Performance | Operator-rated Performance | OPAS Performance |
|---------------------------|----------------------------|-----------------|
| Observer-rated Situation Awareness | 0.79 (0.06)* | 0.05 (0.07) | 0.44 (0.08)* |
| Observer-rated Workload | 0.02 (0.06) | 0.01 (0.07) | -0.08 (0.08) |
| Operator-rated Situation Awareness | - 0.12 (0.07) | 0.48 (0.10)* | 0.07 (0.11) |
| Operator-rated Workload | - 0.06 (0.08) | - 0.22 (0.11) | - 0.09 (0.12) |
| Objective Situation Awareness (SACRI) | 0.11 (0.09) | - 0.14 (0.12) | - 0.05 (0.13) |

4. Discussion and conclusions

From the presented results it was possible to show that overall ratings of workload and situation awareness remained fairly stable in three independent evaluation moments for the same plant. The results also showed a trend to improved performance in the later evaluation.

The data from this study suggests that situation awareness can be a strong predictor of performance, both within the observer-rated measures and within operator-rated measures, confirming the assumption that situation awareness is a relevant aspect for performance in complex systems. The objective measurement of situation awareness did not predict any of the task performance results. This suggests that the type of ratings that are used (observer-rated, operator-rated, or objective) seems to be determinant on the predictability of the measure.

The assumption that workload is an important factor for situation awareness was partially confirmed in the current analysis, since workload and situation awareness measurements appeared to correlate within observer and operator-rated measures.

Overall the correlations tended to be more common and/or stronger within raters than between raters measures, and this aspect was visible in two ways: 1) for example, the observer-rated performance measures were predicted by observer-rated situation awareness measures, but not by any other situation awareness measure; and 2) for example, the observer-rated and operator-rated measures for situation awareness did not correlate in the global analysis, suggesting that the two groups of raters might have been assessing different aspects.

The final results table (Table 4) reveals a low number of significant workload and situation awareness predictors for all performance measures. A relevant aspect that should be mentioned is the fact that, as mentioned in the introduction, control room performance is dependent of various aspects, from the plant systems to the human-machine interfaces in the control room, going through social and team aspects, and also other cognitive aspects related with decision-making or even basic perception – this implies that the amount of variability that could potentially be explained by workload and situation awareness concepts is, by definition, limited. Nonetheless, our goal with this work was to explore the amount of variability in performance that could actually be explained by
these concepts, in an effort to understand whether developing methods and tools to address them could improve our understanding of human performance.

Overall, based on the results of this exploratory analysis, we can conclude that situation awareness is still a valuable construct to the understanding and prediction of human performance. However, the way the concept of situation awareness is defined and measured, is fundamental, since correlation and prediction in this analysis did differed substantially between the types of raters in the situation awareness and task performance measures.

Some essential aspects to attend to in the future rely on the further exploration of the role of workload in human performance in complex systems, trying to link the workload measurements to situation awareness and overall performance. Simultaneously, one aspect that we evaluate as crucial is the understanding of how the type/nature of the measure, namely the type of raters that are selected, can impact the results of the evaluation and also the constructions of models of human performance in the specific context of the control room. Ultimately, being able to select the most reliable and valid sources for information (whether they are self-rating, process expert knowledge, direct observations, objective measures, or any combination), would contribute to a more efficient and robust practice in control room evaluations, hopefully reducing the amount of measures and time that are required to achieve a weighted decision on whether the control room supports safe operation.

References

[1] J.M. O’Hara, J.C. Higgins, S.A. Fleger, & P.A. Pieringer. Human Factors Engineering Program Review Model (NUREG-0711, Rev. 3). (2012). USA: Nuclear Regulatory Comission.
[2] W. F. Stubler, J.M. O’Hara, J.C. Higgins, & J. Kramer. Human-System Interface and Plant Modernization Process: Technical Basis and Human Factors Review Guidance (NUREG/CR-6637). (2004). USA: Nuclear Regulatory Comission
[3] M. R. Endsley, Situation Awareness Misconceptions and Misunderstandings. Journal of Cognitive Engineering and Decision Making, 9, (2015) 4-32, doi:10.1177/155343415572631
[4] P. Berggren, E. Prytz, B. Johansson, & S. Nahlinder. Development of Human Performance Measures for Human Factors Validation in the Advanced MCR of APR-1400. IEEE Transactions on Nuclear Science, 54(6), (2007) 2687–2700. doi:10.1109/TNS.2007.907549
[5] J. S. Ha, P. H. Seong, M. S. Lee, & J.H. Hong. Development of Human Performance Measures for Human Factors Validation in the Advanced MCR of APR-1400. IEEE Transactions on Nuclear Science, 54(6), (2007) 2687–2700. doi:10.1109/TNS.2007.907549
[6] P. Berggren, E. Prytz, B. Johansson, & S. Nahlinder. The relationship between Workload, Teamwork, Situation Awareness, and Performance in Teams: A microworld study. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55(1), (2011). 851–855. doi:10.1177/1071181311551177
[7] M. R. Endsley, Toward a theory of situation awareness in dynamic systems. Human Factors: the Journal of the Human Factors and Ergonomics Society, 37(1). (1995) 32–64. [1] P. Berggren, E. Prytz, B. Johansson, & S. Nahlinder. The relationship between Workload, Teamwork, Situation Awareness, and Performance in Teams: A microworld study. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55(1), (2011). 851–855. doi:10.1177/1071181311551177
[8] C. Khrel, & N. Balf. Cognitive Workload Analysis in Rail Signaling Environments. Cognition, Technology, and Work, 16(3), (2013), DOI: 10.1007/s10111-013-0266-7
[9] B. Sætrevik, & J. Eid. The “Similarity Index” as an Indicator of Shared Mental Models and Situation Awareness in Field Studies. Journal of Cognitive Engineering and Decision Making, 8(2), (2014). 119–136.
[10] R. Ma, & B. Kaber. Situation Awareness and Workload in Driving while Using Adaptative Cruise Control and a Cell Phone. International Journal of Industrial Ergonomics, 35, (2005), 939-953.
[11] S. Nazir, S. Colombo, & D. Manca. The role of Situation Awareness for the Operators of Process Industry. Chemical Engineering Transacions, 26, (2012), 303 – 307.
[12] N. Lau, G.A. Jamiensson, G. Skraaning Jr. Situation Awareness in Process Control: A Fresh Look. NPIC&HMIT Proceedings. (2012). USA: San Diego, 22-26 July.
[13] G.A. Miller. What is Information measurement? In W.C. Howell and I.L. Goldstein (Eds), Engineering Psychology – Current Perspectives in Research. (1971), 132-142. New York, USA: Appleton-Century-Crofts
[14] S.G. Hart, & L. Staveland. Development of the NASA task load index (TLX): Results of empirical and theoretical research. In P. A. Hancock and N. Meshikati (Eds.), Human mental workload, (1988). pp139-183. Amsterdam: North-Holland
[15] S. Dekker, & E. Hollnagel. Human factors and folk models. Cognition, Technology & Work, 6(2), (2004), 79–86. doi:10.1007/s10111-003-0136-9.
[16] K. C. Hendy. Situation Awareness and Workload: Birds of a Feather? AGARD AMP Symposium on Situational Awareness: Limitations and Enhancements in the Aviation Environment. (1995). Brussels, 24-28 April
[17] D. Hogg, K. Follesø, B. Torralba, & F.S. Volden. Measurement of the operator’s situation awareness for use within process control research: Four methodological studies. (HWR-377). (1994). Halden, Norway: OECD Halden Reactor Project.
[18] G. Skraaning, Jr. The operator performance assessment system (OPAS). (HWR-538). (1998). Halden, Norway: OECD Halden Reactor Project.