Human Factors and Simulation, Vol. 83, 2023, 89–98
https://doi.org/10.54941/ahfe1003565

Behavioral Indicators – An Approach for Assessing Control Room Operators' Excessive Cognitive Workload?

Per Øivind Braarud and Giovanni Pignoni

Institute for Energy Technology/NEA Halden Human Technology Organisation Project, Norway

ABSTRACT

Cognitive workload that deteriorates the control room team's performance is a central topic for human-technology design and evaluation. However, while stated as an essential research topic, the literature provides few studies investigating the excessive cognitive workload of complex dynamic human-system work. Multiple techniques have been developed to sample workload. Still, they all struggle to determine the nature of excessive workload, capturing change but leaving the interpretation to the investigator. To advance the measurement of excessive cognitive workload of complex work, this paper proposes to investigate behavioral indicators. Behavioral-based methods differ from performance measures as they concentrate on the operator's behavior rather than the outcome of the actions. The information embedded in the operator's behavior may not directly reflect the outcome of the task. The paper proposes indicator categories in terms of task prioritization, work practices and low-level behavior. The approach implies developing an understanding of how control room teams adapt to and manage task load and how operators are affected by high workload – for the identification of indicators, and for the development and validation of measures from these cognitive workload indicators. The paper presents an initial review of simulator studies identifying adaption such as down-prioritizing secondary tasks, reducing attention to global process overview, asking for or providing team support on task demand, reducing verification of work, and delayed response in communication. Furthermore, we briefly consider the technical and staffing requirements necessary to support these measures.

Keywords: Cognitive workload, Excessive cognitive workload, Behavioral indicators, Control room operators, Workload redline

INTRODUCTION

Importance of Overload

Operator performance is what human factors engineering is ultimately concerned with (Lysaght et al., 1989; O'Donnell and Eggemeier, 1986). Of particular interest is workload that negatively impacts operator behavior and potentially degrades system performance. While both too low and too high cognitive workload can lead to sub-optimal performance, high and excessive workload is the main concern for nuclear control rooms (Lysaght et al., 1989; Wickens et al., 2012). When workload increases, the operator can enhance effort and maintain performance if cognitive capacity is not exceeded (Lysaght et al., 1989; Moray, 1988). However, this suggests there is a point at which the cognitive workload becomes too high for the operator to maintain acceptable performance. Researchers have discussed this topic under the labels of a workload ‘redline’ or a ‘redline zone’ (Colle and Reid, 2005; Hart and Wickens, 1990; Young et al., 2015). Professional operators are trained at managing high workload, for example, by performance regulation regarding the prioritization of tasks and by adapting the division of work within the team (Hockey, 1997; Moray, 1988; Vicente, 1997). Consequently, there may be a zone of substantial-high workload manageable by the operators and with limited degradation of performance. For human-system tests and evaluation, it is critical to identify if cognitive workload approaches or enters a redline and whether workload management is resilient to high workload. The implications of high but manageable cognitive workload observed in evaluation should be well understood before deciding on the human-system acceptability for actual operation. Unfortunately, approaches for assessing redline zones and acceptable cognitive workload are lacking. Researchers have stated this an essential research topic and practitioners frequently request techniques for measuring and assessing workload redline zones (Wickens, 2017; Young et al., 2015).

Managing and Adapting to High Workload in Professional Settings

Skilled workers adapt to high task demand, and performance decrements in work settings are frequently modest in magnitude (Hancock and Warm, 1989; Hockey, 1997). In very high workload conditions, performance is affected, but borderline regions to excessive workload can be characterized by performance protection, in which operators invest effort and work hard (de Waard and Lewis-Evans, 2014; Hockey, 1997). Consequently, a high workload does not necessarily manifest itself in easily observable primary performance outcomes. However high workload might be observable through operators’ workload management and adaption to the situation. The operators’ self-regulation includes monitoring task demands and strategies for coping with demand and modifying their actions to satisfy system goals (Hancock and Matthews, 2019; Hockey, 1997). This might involve modifying performance criteria, such as aiming for sufficient rather than perfect goal achievement. Furthermore, less important tasks might be down prioritized and work practices might be adapted to focus on work on the critical outcome. For example, generally robust but resource demanding, work routines might be sacrificed to maintain critical task performance.

Examples of Behavioral Indicators in Relation to Workload

Researchers have utilized behavioral indicators for several workload-related classifications. Indicators include how the pilot prioritizes tasks and verifies task completion, time management, offering/accepting assistance and help, delegating tasks, and recovery from disruptions (Rondon and Fontes, 2017).

Coppin notes how physiological indexes such as eye behavior can be integrated with behavioral indicators for the categorization of cognitive workload (Coppin, 2019). Furthermore, Schulte and Donath discuss operator behavioral patterns related to self-regulation mechanisms for the adaption to task and workload demands (Schulte and Donath, 2011), Kim et al. proposed that the frequencies of operators' cognitive and communicative activities could be used as indicators of cognitive workload (Kim et al., 2014), and Chung et al. proposed that patterns of the control room team's communication threads such as simultaneous threads and delayed response could serve as workload indicators (Chung et al., 2009). A simulator study with the participation of professional control room operators, found that operators speech features and their human-system interaction activities could be used to classify self-reported cognitive workload with an accuracy of .72 (Braarud et al., 2021). The most important speech features were pitch, amplitude, and articulation rate, and the most important human-system interaction were the number of alarms and frequency of operator–system interaction. Regarding alarms and management of workload, Bliss and Dunn noted how operators experiencing high workload tended to disregard alarms (Bliss and Dunn, 2000).

Workload Measures

Workload measurements are commonly classified into three categories: self-report, performance-based, or physiological (Moray, 1988; Rusnock and Borghetti, 2018). Additionally, we would like to consider a fourth, identified as operator activity or behavioral (Chen et al., 2013). Self-report measures have been popular due to ease of use and low cost of application (Megaw, 2005). However, the validity of self-report measures has been questioned (de Winter, 2014; Matthews et al., 2020). Performance-based measures rely on the notion of limited resources (Yeh and Wickens, 1988). In *Primary task* measures, as the demands of the primary task increase, performance is expected to deteriorate, eventually to the point that it does not satisfy the safety criteria. In *Secondary task measures* the performance of a secondary task is expected to capture the operator's remaining capacity while performing a primary task (Mulder, 1979; Tsang and Vidulich, 2006). Physiological approaches attempt to measure cognitive workload processes through physiological phenomenon reflecting the underlying activity of the central nervous system. The most common physiological measurement concerns heart activity, brain activity, respiratory rate, galvanic skin response, and ocular behavior (Charles and Nixon, 2019; Megaw, 2005).

Behavioral-based methods fundamentally differ from performance measures as they concentrate on the operator's behavior rather than the outcome of the actions, the information embedded in the operator's behavior is not directly reflected in the outcome of the task (Chen et al., 2013). Response-based behavioral features characterize deliberate/voluntary user activity and are therefore differentiated from physiological measures that concentrate on involuntary responses. These include activities that are part of the naturally occurring task, such as interface navigation (eye tracking, mouse or keyboard capture), verbal communication, and multiple speech-related parameters

(Chen et al., 2013). These are objective metrics that can be collected while the participant is performing the task without disrupting it and consequently, measurement is usually non-intrusive. The measurement assumes that patterns of operator behavior are affected by one's cognitive state, including workload. Two types of information can be extracted: the response itself and its inherent meaning in the context of the task (i.e. searching for an item in an interface) as well as the manner of execution (a thorough search compared to a frantic one will be characterized by different visual attention patterns, different mouse movement dynamics, etc.) (Chen et al., 2013). Figure 1 is an attempt to illustrate the relation between physiological measures, behavioral indicators, and task performance measures. Low-level behaviors are often countable or quantifiable, often unconscious, abstractions from the deliberate task work and work practices. These partially overlap with commonly used physiological indicators of workload. Work practices and task prioritization include high-level behaviors that may require qualitative evaluation with the support of a subject matter expert. Work practices and task prioritization underlie task performance across a variety of situations, both in foreseen and unforeseen situations, but are not a direct measure of performance.

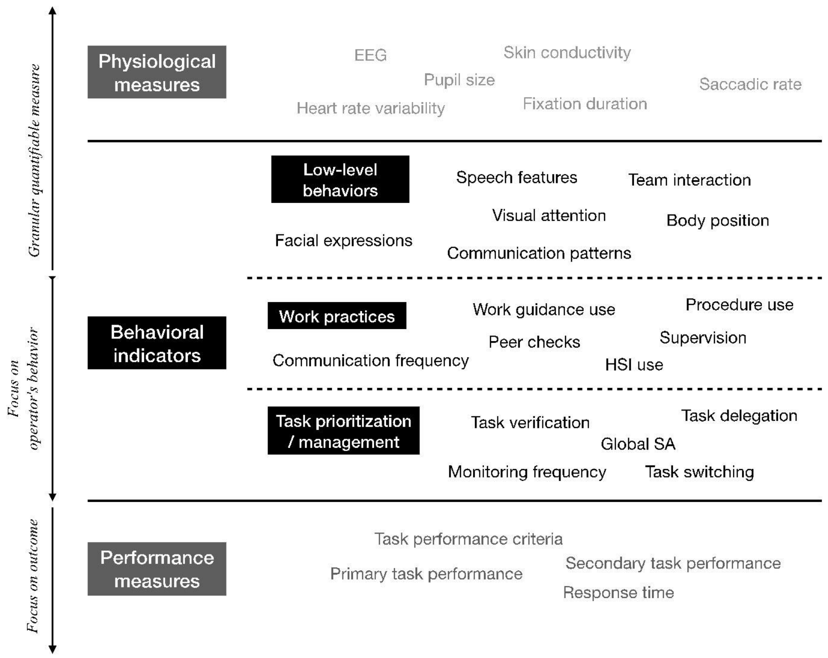

Figure 1: The figure illustrates the relation between behavioral, physiological, and task performance measures.

BEHAVIOURAL INDICATORS FROM HALDEN SIMULATOR STUDIES

To identify potential behavioral indicators of operator adaption to high or excessive cognitive workload we reviewed simulator studies with professional nuclear operators from the Halden Human-Machine Laboratory. We selected reports ranging from 2001 to 2020. The review focused on task prioritization and work practices. All technical reports were initially reviewed and reports describing scenarios as knowledge-based, beyond design basis, unexpected events, un-familiar events, limited guided by operating procedures, or frankly describing the scenario(s) as complex or demanding for the operators were selected. In addition, reports including topics on work practices in the control room were selected. This resulted in twenty-one technical reports eligible for our analysis. Table 1 provides examples identified by the review of potential behavioral indicators related to management of and adaption to cognitive workload.

EXCESSIVE WORKLOAD AND BEHAVIORAL INDICATORS

From the literature and the Halden simulator studies, we did not identify studies that explicitly investigated behavioral indicators across cognitive workload levels and explicitly addressed the measurement of excessive cognitive workload. But our hunch is that many behavioral indicators change nature when load increases from smoothly manageable to manageable by extensive adaption. Figure 2 provides a simple illustration of hypothesized patterns of behavioral indicators across increasing level of cognitive workload. If correct, this hypothesis has strong implications for measurement utilizing behavioral indicators. A shift in the nature of the indicators might be a good indicator of excessive cognitive workload.

We hypothesize the pattern of key indicators will show a significant deviation if the operator experiences excessive cognitive workload. The trend will vary between different indicators, and it should be expected that they

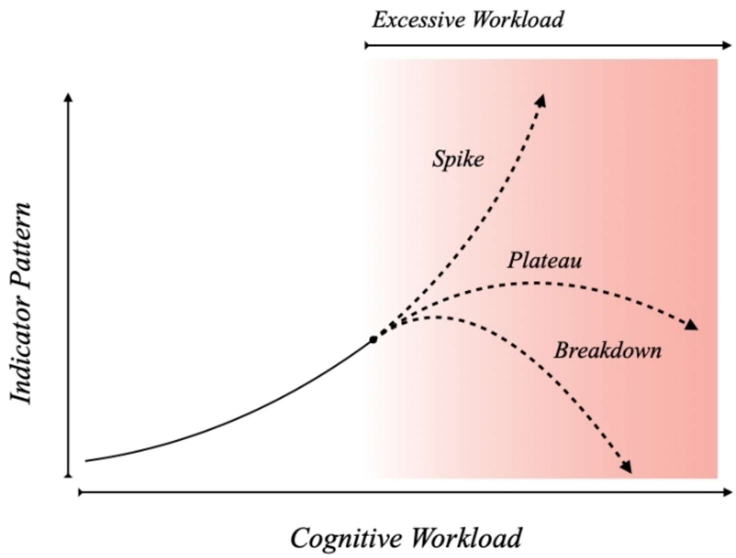

Figure 2: Hypothesized behavioral indicators patterns in relation to excessive workload.

Table 1. Behavior indicators on high cognitive workload from Halden simulator studies.

Behavioral indicator	Example
Reduced attention to global plant overview.	In demanding scenarios with misleading and failed information, operators' overall plant overview was reduced (Nystad et al., 2020a, 2020b). A study of staffing levels found that maintaining an overall overview of the plant status and automation while dealing with disturbances and system failures may have exceeded the capacity of the operators (Eitrheim et al., 2010).
Utilization of non-critical HSI deteriorate.	A computerized system for Technical Specification worked well for normal operation but was challenging to use in time-pressured scenarios with high information load from the support system (Braarud and Svengren, 2020).
Acknowledgement of alarms deteriorates	A study observed that operators altered their alarm management (alarm acknowledgement suffered) in complex high load scenarios compared to simpler scenarios. (Broberg et al., 2018).
Increased navigation to verify or complement uncertain information.	In scenarios with misleading and failed information, operators navigated to check alternative info, e.g., comparing channels, checking redundant information (Nystad et al., 2020b).
Stuck in an operating procedure step.	A study of accident scenarios observed crews stopping in procedure steps for 10–15 minutes presumably due challenge in understanding the plant response. The analysis suggested high cognitive workload for this period of work (Massaiu and Holmgren, 2017).
Distributing workload within the team	A study reported that the supervisor took over operation tasks from the operators in scenarios of high workload. For example, monitoring trends, overviewing alarm status, and procedure foldout pages (Massaiu and Holmgren, 2017). A study on teamwork found that control room operators allocated task to the supervisor when they experienced high load (Skjerve et al., 2008). Furthermore, in cases of an overloaded supervisor, the reactor operator assisted on team leadership (Massaiu and Holmgren, 2017).
"Break down" or deterioration of communication threads	In a scenario with many degraded plant indications, crews questioned and started to discuss the information but did not conclude or complete assessment of the status due to becoming occupied with other tasks. Examples of the last step of three-way communication deteriorated in highly demanding scenarios (Nystad et al., 2020a).

may change with different direction and rate. For example, when the cognitive workload in the control room increases, we would expect the level and formality of communication and team discussions to increase. If the workload becomes excessive the level of communication and correct communication procedure might be considerably reduced and appear scattered.

Similarly, the use of supporting human-system interfaces might increase and be utilized according to established work practices until excessive workload causes reduced and unsystematic utilization.

FEASIBILITY AND REQUIREMENTS FOR SIMULATOR STUDIES

The use of behavioral indicators in simulator studies requires few technical and staffing resources. Technically, it may suffice to have a mean to record audio and video for post-facto coding and have a good overview of the teams' actions. More specific tools might be needed to automatically track and log indicators such as gaze or movement in space and to extract speech features from communication. Staffing requirements will include one or more process experts to evaluate the relevant indexes and develop a coding system specific to the scenario and task in question. Heyman et al. describe basic forms of behavioral observation coding systems: *Topographical* coding systems, which measure the occurrence of behaviors and *Dimensional* coding systems measure the intensity of behaviors (Heyman et al., 2014). Ideally, it could be possible to find an existing system that can be adapted to the specific needs (scenario, type of plant, etc.). In this regard, Heyman et al. remind that one should establish the validity of a coding scheme in its context. In our case, the coding scheme itself represents a hypothesis about the implications of cognitive workload for operator behavior. Therefore, a step prior to data collection involves a researcher together with a process expert, evaluating the task and scenario to identify relevant behaviors and how to quantify them. Literature can inform and narrow this choice, learning which coding systems were used or which behaviors have been observed to change in function of workload manipulations. Content validity, (whether the measurement items are relevant and representative of nuclear operators) will be dependent on the theoretical and analytical basis of the indicators.

CONCLUSION

Behavioral indicators are being used for investigating a number of human performance topics. However, there is limited utilization of this type of indicators for understanding cognitive workload of control room work. We hypothesize that behavioral indicators can be utilized to develop specific workload measurement with improved content validity, improved sensitivity over performance measures, and improved temporal resolution compared to commonly applied self-report measures. The combined use of behavioral indicators with more traditional approaches has the potential to support relevant activities such as human factors engineering validation, workload method development and verification, as well as improved understanding of workload dynamics and excessive cognitive workload. Behavioral indicators can help point to aspects of human-system design related to excessive load e.g., task priority, work practices and communication issues. This approach will contribute to our understanding of causes of primary task breakdown or near breakdown, provide increased utility for the design of human-system interfaces, and provide knowledge for training on workload management.

In summary, we propose that behavioral indicators represent a practical and useful approach for investigating high and excessive cognitive workload in complex dynamic work settings - addressing a critical gap in the literature. Future research could improve and expand the definition of behavioral indicators and clarify their relationship with physiological and performance indicators. Furthermore, approaches for the identification of operational definitions of indicators and their use in the analysis of operator cognitive workload are needed. The validity of this type of approach as well as the utility compared with other approaches to cognitive workload assessment should be investigated.

ACKNOWLEDGMENT

The technical reports reviewed in this paper were founded by the OECD Halden Reactor Project, Norway.

REFERENCES

Bliss, J. P., Dunn, M. C., 2000. Behavioural implications of alarm mistrust as a function of task workload. Ergonomics 43, 1283–1300.

Braarud, P. Ø., Bodal, T., Hulsund, J. E., Louka, M. N., Nihlwing, C., Nystad, E., Svengren, H., Wingstedt, E., 2021. An Investigation of Speech Features, Plant System Alarms, and Operator–System Interaction for the Classification of Operator Cognitive Workload During Dynamic Work. Hum. Factors J. Hum. Factors Ergon. Soc. 63, 736–756. https://doi.org/10.1177/0018720820961730

Braarud, P. Ø., Svengren, H., 2020. Evaluation of Plant Overview Procedures Comparing computerized and paper-based solutions (No. HWR-1299). OECD Halden Reactor Project, Halden, Norway.

Broberg, H., Hildebrandt, M., Nowell, R., 2018. Results from the 2010 HRA Data Collection at a US PWR Training Simulator (No. HWR-981). OECD Halden Reactor Project, Halden, Norway.

Charles, R. L., Nixon, J., 2019. Measuring mental workload using physiological measures: A systematic review. Appl. Ergon. 74, 221–232.

Chen, F., Ruiz, N., Choi, E., Epps, J., Khawaja, M. A., Taib, R., Yin, B., Wang, Y., 2013. Multimodal behavior and interaction as indicators of cognitive load. ACM Trans. Interact. Intell. Syst. TiiS 2, 1–36.

Chung, Y. H., Yoon, W. C., Min, D., 2009. A model-based framework for the analysis of team communication in nuclear power plants. Reliab. Eng. Syst. Saf. 94, 1030–1040. https://doi.org/10.1016/j.ress.2008.11.010

Colle, H. A., Reid, G. B., 2005. Estimating a Mental Workload Redline in a Simulated Air-to-Ground Combat Mission. Int. J. Aviat. Psychol. 15, 303–319. https://doi.org/10.1207/s15327108ijap1504_1

Coppin, G., 2019. Operator Functional State: Measure It with Attention Intensity and Selectivity, Explain It with Cognitive Control, in: Human Mental Workload: Models and Applications: Third International Symposium, H-WORKLOAD 2019, Rome, Italy, November 14–15, 2019, Proceedings. Springer, p. 156.

de Waard, D., Lewis-Evans, B., 2014. Self-report scales alone cannot capture mental workload: A reply to De Winter, Controversy in human factors constructs and the explosive use of the NASA TLX: a measurement perspective. Cogn. Technol. Work 16, 303–305. https://doi.org/10.1007/s10111-014-0277-z de Winter, J. C. F., 2014. Controversy in human factors constructs and the explosive use of the NASA-TLX: a measurement perspective. Cogn. Technol. Work 16, 289–297. https://doi.org/10.1007/s10111-014-0275-1

Eitrheim, M. H. R., Jr., G. S., Lau, N., Karlsson, T., Nilwing, C., Hoffmann, M., Farbrot, J. E., 2010. Staffing Strategies in Highly Automated Future Plants - Results from the 2009 HAMMLAB Experiment (No. HWR-938). OECD Halden Reactor Project, Halden, Norway.

Hancock, P. A., Matthews, G., 2019. Workload and Performance: Associations, Insensitivities, and Dissociations. Hum. Factors J. Hum. Factors Ergon. Soc. 61, 374–392. https://doi.org/10.1177/0018720818809590

Hancock, P. A., Warm, J. L., 1989. A Dynamic Model of Stress and Sustained Attention. Hum. Factors J. Hum. Factors Ergon. Soc. 31, 519–537. https://doi.org/10.1177/001872088903100503

Hart, S. G., Wickens, C. D., 1990. Workload Assessment and Prediction, in: Booher, H. R. (Ed.), Manprint. Springer Netherlands, Dordrecht, pp. 257–296. https://doi.org/10.1007/978-94-009-0437-8_9

Heyman, R. E., Lorber, M. F., Eddy, J. M., West, T. V., 2014. Behavioral Observation and Coding, in: Reis, H. T., Judd, C. M. (Eds.), Handbook of Research Methods in Social and Personality Psychology. Cambridge University Press, pp. 345–372. https://doi.org/10.1017/CBO9780511996481.018

Hockey, G. R. J., 1997. Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. Biol. Psychol. 45, 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4

Kim, S., Kim, Y., Jung, W., 2014. Operator's cognitive, communicative and operative activities based workload measurement of advanced main control room. Ann. Nucl. Energy 72, 120–129. https://doi.org/10.1016/j.anucene.2014.04.019

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., Linton, P. M., 1989. Operator Workload: Comprehensive Review and Evaluation of Operator Workload Methodologies.

Massaiu, S., Holmgren, L., 2017. The 2013 Resilient Procedure Use Study with Swedish Operators: Final Results (No. HWR-1216). OECD Halden Reactor Project, Halden, Norway.

Matthews, G., De Winter, J., Hancock, P. A., 2020. What do subjective workload scales really measure? Operational and representational solutions to divergence of workload measures. Theor. Issues Ergon. Sci. 21, 369–396. https://doi.org/10.1080/1463922X.2018.1547459

Megaw, T., 2005. The definition and measurement of mental workload. Eval. Hum. Work 3, 541–542.

Moray, N., 1988. Mental workload since 1979. Int. Rev. Ergon. 2, 123–150.

Mulder, G., 1979. Mental load, mental effort and attention, in: Mental Workload. Springer, pp. 299–325.

Nystad, E., Kaarstad, M., McDonald, R., 2020a. Decision-making with degraded HSI process information (No. HWR-1242). OECD Halden Reactor Project, Halden, Norway.

Nystad, E., Kaarstad, M., Nihlwingand, C., McDonald, R., 2020b. Functionalities to support operation during situations with limited process information (No. HWR-1243v2). OECD Halden Reactor Project, Halden, Norway.

O'Donnell, R. D., Eggemeier, F. T., 1986. Workload assessment methodology, in: Handbook of Perception and Human Performance, Vol. 2. Cognitive Processes and Performance. John Wiley & Sons., pp. 1–49.

Rondon, M., Fontes, R., 2017. Reflections on Automation and the Need for New Competencies in the Civil Pilot Training. Aeron Aero Open Access J 1, 139–150.
Rusnock, C. F., Borghetti, B. J., 2018. Workload profiles: A continuous measure of mental workload. Int. J. Ind. Ergon. 63, 49–64. https://doi.org/10.1016/j.ergon.2016.09.003
Schulte, A., Donath, D., 2011. Measuring Self-adaptive UAV Operators' Load-Shedding Strategies under High Workload, in: Harris, D. (Ed.), Engineering Psychology And Cognitive Ergonomics, Lecture Notes in Artificial Intelligence. pp. 342–351.
Skjerve, A. B., Nihlwing, C., Nystad, E., 2008. Lessons Learned From The Extended Teamwork Study (No. HWR-867). OECD Halden Reactor Project, Halden, Norway.
Tsang, P. S., Vidulich, M. A., 2006. Mental workload and situation awareness.
Vicente, K. J., 1997. Operator adaptation in process control: A three-year research program. Control Eng. Pract. 5, 407–416. https://doi.org/10.1016/S0967-0661(97)00018-X
Wickens, C. D., 2017. Mental Workload: Assessment, Prediction and Consequences, in: Longo, L., Leva, M. C. (Eds.), Human Mental Workload: Models and Applications, Communications in Computer and Information Science. Springer International Publishing, Cham, pp. 18–29. https://doi.org/10.1007/978-3-319-61061-0_2
Wickens, C. D., Hollands, J. G., Banbury, S., Parasuraman, R., 2012. Engineering psychology and human performance, Fourth edition. ed. Pearson, Boston.
Yeh, Y.-Y., Wickens, C. D., 1988. Dissociation of performance and subjective measures of workload. Hum. Factors 30, 111–120.
Young, M. S., Brookhuis, K. A., Wickens, C. D., Hancock, P. A., 2015. State of science: mental workload in ergonomics. Ergonomics 58, 1–17. https://doi.org/10.1080/00140139.2014.956151