A Comparative Analysis of Workload for Navigation Tasks Performed onboard and at Simulated Remote Control Centers for MASS Using NASA-TLX

Ender Asyali[1]*, Stephen Tarrant[1], Adam Slazas[1], Steve Cole[1]

1Maine Maritime Academy, USA
** Corresponding author: ender.asyali@mma.edu Tel.: +1-207-326-2648*

Abstract: The maritime industry is on the verge of dramatic technological change as the maritime world transitions to the development of Maritime Autonomous Surface Ships (MASS). The development of MASS will affect how the maritime workforce is trained and how it operates. Remote Control Centers (RCC) for MASS will be the main work environment during the MASS era. Many potential human factors, technical, and regulatory challenges have already been analyzed. In this study, a comparative analysis was performed between real life navigation task workload and simulated RCC navigation task workload. NASA-TLX (Task Load Index) was used to measure the task load of navigators while underway compared to a similar simulated voyage. Results showed that there are significant differences in perception of workload between onboard and both simulated monitor and mockup-based workplaces with the largest difference consisting of the mental demand of being on a bridge versus being in a monitor or mockup workplace. Topics for future studies include layouts of future RCCs and engagement of operators.

Keywords: Workload, Remote Control Center, MASS, NASA-TLX

1. Introduction

Maritime education and training (MET) institutions are important stakeholders in the maritime transportation system. MET institutions must prepare, equip, and support students to successfully adopt the technology needed to operate and manage MASS ships in the coming highly competitive and changing workplace. RCCs will be the main work environment during the MASS era. The transition of deck officers from ship's bridges to RCCs will change important human performance factors including workload, stress, and fatigue (Wahlström et al. 2015), and the state of operator situational awareness (Porathe et al. 2014; Man et al. 2015; Mackinnon et al. 2015). Additionally, the uncertainty related to both the human-machine and human-human interactions will affect operations (Kari and Steinert 2021), operator's stress in different levels of workload (Kari et al. 2019), ship-sense and harmony (Man et al. 2014), design criteria for the Human-Machine Interface (HMI), safe and secure transfer of very large data quantities (Porathe 2014), operational, regulatory, and quality challenges (Komianos 2018), cybers risks (Andersen 2018), and autonomous system design to meet the STCW requirements (Dittmann et al. 2021).

1.1 Remote *Control Centers*

As MASS develops, there will be a need to create shoreside RCCs where a human will be able to remotely monitor and/or operate one or more vessels simultaneously, and where some or all vessel functions can be executed (AUTOSHIP 2023). An RCC, also called Remote Operation Center, implies that responsibilities related to planning operations, maintenance activities, and logistics can be included (AUTOSHIP 2023). The RCC should allow for a remote operator to monitor, and if necessary, to take control of all operational functions of the vessel including navigation, VHF communications, and ship's systems via a real-time, reliable communications system. The IMO defines four degrees of autonomy for ships, ranging from Levels 1 and 2, where a seafarer is present on the ship and RCC is used to assist in decisions (Level 1) or take over if necessary (Level 2) to Levels 3 and 4 where no seafarer is present and the ship is operated by remote control (Level 3) or is fully autonomous (Level 4) (Table 1).

Table 1. Levels of autonomy and RCC requirements for MASS as defined by IMO.

LoA (Level of Autonomy)	**Seafarer on board**	**Function of seafarer**	**Who makes the decisions and takes actions**	**RCC required**
1-Ship with automated processes and decision support	yes	to operate and control / ready to take control of shipboard systems and functions	Seafarer	no
2-Remotely controlled ship	yes	available to take control and to operate the shipboard systems and functions	RCC and/or Seafarer	yes
3-Remotely controlled ship	no.	N/A	RCC	yes
4-Fully autonomous ship	no	N/A	The operating system of the ship	yes, implied

Source: Based on IMO, 2021

1.2 Workload

Workload is an essential factor to consider for any task including standing watch in an RCC. Lysaght et al. (1989) defined workload as the relative capacity to respond. Hart and Staveland (1988) describe workload as the perceived relationship between the amount of mental processing capability or resources and the amount required by the task. According to Hoonakker et al. (2011), workload is a complicated relationship between operator, external physical or cognitive demand, and performing a certain task. Other factors, including environmental, organizational, and psychological also play a role in this process.

There are many examples of accidents related to both high and low levels of workload in the shipping industry. High workload can overwhelm a watchstander, reducing performance and leading to accidents. Low workload can cause inattentiveness and complacency as demonstrated by accidents that have occurred in good visibility in daytime and in open water with low traffic conditions. Simply said: too little workload can be just as detrimental as too great a workload (Koester, 2002; Lysaght et al, 1989). What is needed in an RCC is to create an environment where an operator has a moderate workload that will keep them engaged and attentive while not being overloaded. This has implications on how workstations are designed and how many vessels an operator should be responsible for simultaneously.

2. Research Methodology

2.1 Objectives of the study

In this study a comparative analysis was performed to determine the navigation task load of watchstanders between real life and two different settings of simulated RCC workspaces by using NASA-TLX. The monitor based RCC workplace was made up of eight monitors; and the mockup RCC was a fully simulated ship's bridge with ECDIS, radar, VHF and other hands-on ship controls. Although there are other tasks that were performed by cadet watchstanders, the focus of this study is on the task of navigation which consists of route monitoring, execution of a passage plan, keeping proper lookout, maintaining internal and external situational awareness, checking the performance of navigational equipment, surveillance of the ship, and collision avoidance in compliance with the COLREGS. Research into workload is necessary as part of the design development of RCCs to guide the physical design of the workspace.

2.2 Data collection

NASA-TLX is a widely used, subjective measurement tool used to assess workload on operator(s) in environments with various human-machine interface systems such as aircraft cockpits; command, control, and communication workstations; supervisory and process control; and simulations and laboratory tests (NASA 1986). NASA-TLX compares six factors; namely the mental demand, physical demand, temporal demand, performance, effort, and frustration levels (Hart and Staveland 1988). The literature shows that the NASA- TLX is the most cited, reliable and valid survey-based workload measurement instrument (Sugarindra et al. 2017; Grier 2015; Bjørneseth et al. 2012; Hoonakker et al. 2011).

The study was comprised of two phases. In the first phase, real life navigation tasks were performed by eight cadets aboard the T/S *State of Maine* during the summer training cruises of 2019 and 2022. A voluntary response sampling method was applied. The same cadets took part in the second phase of the study at Maine Maritime Academy's Bridge Navigation Simulator Center in two types of simulated RCC workplace environments (mockup and monitor). The onboard ECDIS recordings and screen shots were saved; and similar area, traffic, visibility, time and date, and sea state conditions were reproduced in Wärtsila, NTPRO 5000 navigation simulator. The cadets in this study had each finished their first-year training cruise, their sophomore cadet shipping experience, and were currently performing their junior training cruise, before entering their senior academic year. According to the Maine Maritime Academy (MMA) Institutional Research Board (IRB) requirements, all participating students were informed about the research and then asked to sign a consent form. The paper and pencil version of NASA-TLX was completed immediately after the completion of the navigation task.

2.2.1 NASA-TLX application process

The experimental procedure of NASA-TLX employs two steps in the evaluation process including source of loads (weights) and magnitude of loads (ratings). Cadets were asked to read the scale definitions and instructions of mental demand (MD), physical demand (PD), temporal demand (TD) performance (PO), effort (EF) and frustration level (FL) from the NASA-TLX instructions. Cadets then performed navigational tasks onboard the T/S *State of Maine* and then again in two different simulated RCC workplace settings.

Determining the source of loads (weights) was performed by completing the NASA-TLX "Sources of Workload Comparison Card". In this stage a pair-wise comparison was performed. Six factors contribute to the workload for a specific task at different weights. These contributions (weights) are calculated by pair-wise comparison among six factors (e.g., mental demand or effort) by circling one factor which contributes more to the workload of the task that was performed. Then the number of times that each factor was selected was tallied and the weight of each factor was calculated. The results were compiled using a source of workload tally sheet. Determining the magnitude of loads (ratings) consisted of completing a rating sheet. Cadets were asked to mark the rating scale sheet for each factor to get the magnitude of that factor for the task they performed. The overall workload score for each cadet is calculated by multiplying each "rating" by the "weight" given to that factor by that cadet. Then the sum of the weighted ratings for each task is divided by 15 which is the sum of the weights. Interpretation of the score of NASA-TLX results are performed as low (0-9); medium (10-29); rather high (30-49); high (50-79); very high (80-100) (Sugarindra et al. 2017).

3. Findings and Analysis

Descriptive statistics about the environmental conditions where navigation tasks were performed revealed that 62.5 % of the areas are open sea, 50 % of the traffic is low, 75% of the time is nighttime, 75 % of the visibility is clear, and 87.5 % of the sea state is less than or equal to Douglas Scale 4 with winds less than or equal to Beaufort Scale 4.

Mean TLX score for navigation tasks performed on board is 60.289 (SD: 11.917). This value is in the range of high workload. Mockup TLX mean score is 32.703 (SD: 14.187) and Monitor TLX mean score is 30.621(SD: 14.923). These correspond to both scoring in the “rather high" range.

Weighed ratings of each workload factor are found for each participant by multiplying source of loads (weights) and magnitude of loads (ratings), then the average weighted ratings were determined. The results for each factor for three different workplaces are shown in Figure 1. Mental demand has the highest weighted rating onboard (M=23.5), where physical demand has the lowest weighted rating for monitor TLX (M=0. 40).

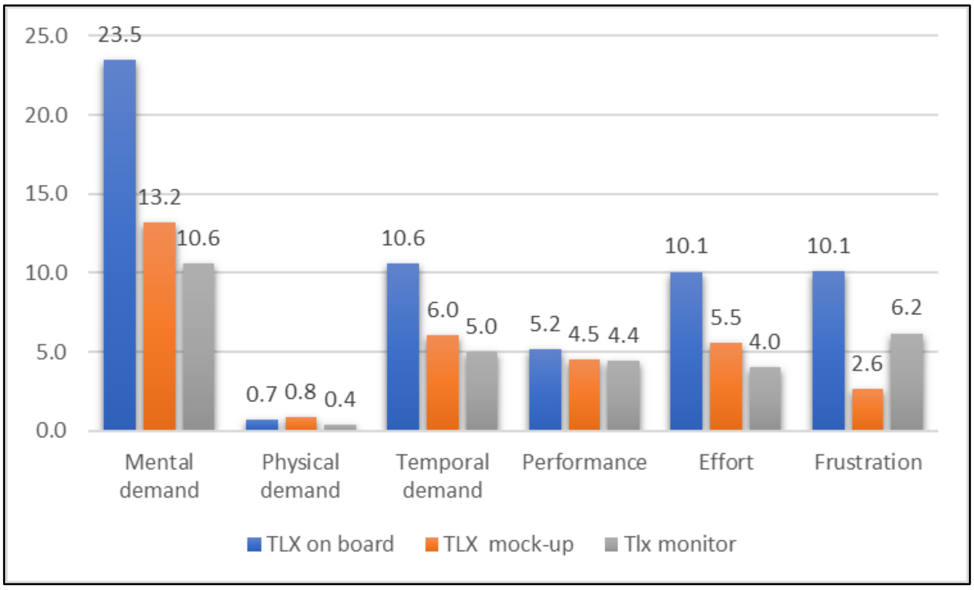

Figure 1: Comparison of weighted ratings of TLX factor scores.

SPSS v:29.0 (Statistical Package for the Social Sciences) program was used for the statistical analysis of the data. Three null hypotheses were developed to determine if there are significant differences between TLX scores of real-life, mockup, and monitor based RCC workplace settings. Hypotheses were tested by paired-samples t-test where scores were taken from the same individuals.

Paired-samples t-test results revealed that there is a significant difference in the scores for onboard TLX and mockup (p = .003). On average, onboard TLX scores were 27.6 points higher than monitor and mockup TLX scores. There is also a significant difference in the scores for onboard TLX and monitor TLX (p = .002). On average, onboard TLX scores were 29.7 points higher than monitor and mockup TLX scores. However, there is not a significant difference in the scores for mockup and monitor TLX conditions (p = .44). On average, mockup TLX scores were 2.1 points higher than monitor TLX scores (Table 2).

Table 2: Results of paired-samples t-test.

Comparison Pairs	**Mean**	**Deviation**	**Error**	**CI lower**	**CI upper**	**t**	**df**	**p**
Onboard/Mockup	27.586	17.463	6.174	12.987	42.186	4.468	7	0.003
Onboard/Monitor	29.667	17.52	6.194	15.02	44.315	4.789	7	0.002
Mockup/Monitor	2.081	7.261	2.567	-3.989	8.152	0.811	7	0.444

4. Conclusion

A comparative analysis of results revealed that there is a significant difference in TLX scores between onboard and those of both simulated RCC mockup and RCC monitor environments. The real-life, on-board workload TLX score is higher than the workload TLX score in both simulated workplace settings. This result is consistent with the earlier findings of Hertzum (2021). Although the mean mockup TLX score is higher than TLX monitor scores, there is no significant difference between RCC mockup TLX scores and RCC monitor TLX scores.

Considering the sub-factors of TLX individually, physical demand was rated similarly across all three workplaces. However, the other TLX sub-factors (mental demand, temporal demand, performance, effort, and frustration) were all rated the highest in the real-life watchstanding environment. A comparison of the mean values of the sub-factors between monitor and mockup alone showed higher values for all sub-factors in the mockup than the monitor except for frustration. Frustration was the only sub-factor higher in the monitor as compared to the mockup. The NASA-TLX manual (1986) defines "Frustration" as "*How insecure, discouraged, irritated, stressed and annoyed versus secure gratified, content relaxed and complacent did you feel during the task*." The mockup environment is more familiar to cadets, but working in front of the monitors is something new to them. This may cause a higher level of frustration and could be the explanation for the higher frustration in the monitor environment.

Narrative statements of participants revealed that they felt more engaged and involved in the mockup than they did in the monitor. As a result, operators may experience "out-of-the-loop" performance problems while performing tasks in monitor-based workplaces. This leaves operators of automated systems handicapped in their ability to take over manual operations in case of automation failure (Endsley and Kiris 1995) which may result in loss of navigational skills, situational awareness, and complacency problems during emergency situations. Since the monitor environment is more likely to be used in future RCCs, more research needs to be done on the impact of frustration on operators as monitor environments provide less involvement and engagement than mockup environments.

5. Future Research

This study opens the discussion about improving the workplace organization, layout and procedures, and competency considerations of operators at MASS RCCs. Further research is needed in developing training requirements and non-technical skills of RCC operators as well as the impact of RCC design on operator workload. Future study should include a larger sample size, more experienced watchstanders monitoring several vessels simultaneously as well as assessing workload during routine vs emergency vessel operations.

References

[1] Andersen J H (2018) Cyber risks of autonomous and remote controlled vessel, DNV GL, Maritime Issues, challenges and opportunities, Conference Artificial intelligence and autonomous ships.

[2] AUTOSHIP (2023) [D8.2] Roadmap for Autonomous ship adoption and development, https://www.autoship-project.eu/news-press/4-new-reports-available/ (Accessed 30 March 2023)

[3] Bjørneseth F B, Renganayagalu S K, Komandur S, Hornecker E (2012) Towards an experimental design framework for evaluation of dynamic workload and situational awareness in safety critical maritime settings September, DOI: 10.14236/ewic/HCI2012.43, Conference: British HCI'12

[4] Dittmann K, Hansen P N, Papageorgiou D, Jensen S, Lützen M, Blanke M (2021) Autonomous surface vessel with remote human on the loop: system design for STCW Compliance,Volume 54, Issue 16, Pages 224-231,

ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2021.10.097.

[5] Endsley M, Kiris E O (1995) The out-of-the-loop performance problem and level of control in automation, The Journal of the Human Factors and Ergonomics Society · June

[6] Grier R A (2015) How High Is High? A meta-analysis of NASA-TLX, global workload scores, Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting

[7] Hart S G, Staveland L E (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research, Ed(s): Peter A. Hancock, Najmedin Meshkati, Advances in Psychology, North-Holland, Vol. 52.

[8] Hertzum M (2021) Reference Values and Subscale Patterns for the Task Load Index (TLX): A Meta-Analytic Review Ergonomics, vol. 64, no. 7 (2021), pp. 869-878 Author version. Published version available at: https://doi.org/10.1080/00140139.2021.1876927

[9] Hoonakker P, Carayon P, et al (2011) Measuring workload of ICU nurses with a questionnaire survey: The NASA Task Load Index (TLX). IIE Trans Health Syst Eng. 2011;1(2):131-143. doi: 10.1080/19488300.2011.609524. Epub 2011 Oct 12. PMID: 22773941; PMCID: PMC3388621

[10] International Maritime Organization (2021) MSC.1/Circ.1638, 3 June 2021 Outcome of The Regulatory Scoping Exercise for the Use of Maritime Autonomous Surface Ships (MASS)

[11] Kari R, Steinert M (2021) Human factor issues in remote ship operations: lessons learned by studying different domains. J. Mar.Sci. Eng. 2021, 9, 385. https:// doi.org/10.3390/jmse9040385

[12] Kari R, Steinert M, Gaspar H. M (2019) EEG application for human-centered experiments in remote ship operations, November 2019 Conference: CENTRIC 2019, The Twelfth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services

[13] Koester T (2002) Human factors and everyday routine in the maritime work domain. https://dvikan.no/ntnu-studentserver/reports/HUMAN%20FACTORS%20AND%20EVERYDAY%20ROUTINE%20IN%20the%20maritime%20work%20domain%27.pdf (Accessed February 2023)

[14] Komianos A (2018) The Autonomous Shipping Era. Operational, Regulatory, and Quality Challenges TransNav the International Journal on Marine Navigation and Safety of Sea Transportation 12(2):335-348 DOI: 10.12716/1001.12.02.15

[15] Lysaght R J, Hill S G, Dick A O, Plamondon B D, Linton PM, Wierwille W W. (1989) Operator workload: comprehensive review and evaluation of operator workload methodologies (No ARI Tech Report 851) Fort Bliss, TX: U.S. Army Research Institute

[16] Mackinnon S N, Man Y, Lundh M, Porathe T (2015) Command and control of unmanned vessels: keep shore-based operators in-the-loop, Conference: ATENA Conferences System, NAV 2015 18th International Conference on Ships and Shipping Research

[17] Man Y, Lundh M, Porathe T, MacKinnon S (2015) From desk to field - human factor issues in remote monitoring and controlling of autonomous unmanned vessels, Procedia Manufacturing 00 (2015) 000–000 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

[18] Man Y, Lundh M, Porathe T (2014) Seeking harmony in shore-based unmanned ship handling - from the perspective of human factors, what is the difference we need to focus on from being onboard to onshore? Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014, Kraków, Poland 19-23 July 2014 Edited by T. Ahram, W. Karwowski and T. Marek

[19] NASA (1986) Task Load Index (TLX) v.10 paper and pencil package, Human Performance RESEARCH Group. (https://humansystems.arc.nasa.gov/groups/tlx (Accessed 2 March 2023)

[20] Porathe, T. (2014). Remote monitoring and control of unmanned vessels –The MUNIN Shore Control Centre. Proceedings COMPIT ‘14, 12-14 May 2014 in Redworth, UK. V. Bertram (Ed.) Hamburg, Technische Universität Hamburg- Harburg, 2014, ISBN 978-3-89220-672-9 460

[21] Porathe T, Prison J, and Man Y (2014) Situation awareness in remote control centers for unmanned ships, Human Factors in Ship Design & Operation, The Royal Institution of Naval Architects, 26-27 February 2014, London, UK

[22] Sugarindra M. et al (2017) Mental workload measurement in operator control room using NASATLX, IOP Conf. Ser.: Mater. Sci. Eng. 277 012022 DOI 10.1088/1757-899X/277/1/012022

[23] Wahlström M, Hakulinen J, Karvonen H, Lindborg I (2015) Human factors challenges in unmanned ship operations: insights from other domains, July, Procedia Manufacturing 3:1038–1045, DOI: 10.1016/j.promfg.2015.07.167