An assessment of workload in a simulated submarine control room

Kiome A. Pope, Aaron P.J. Roberts & Neville A. Stanton

University of Southampton, UK

ABSTRACT

The Command Team Experimental Testbed (ComTET) is a programme of work designed to study the operation of current and future submarine command teams. As part of this work a submarine control room simulator was designed and built for testing purposes. Baseline testing was conducted to establish current functionalities, with recommendations forming the basis for Manipulation One where some operators were co-located. During testing, 32 participants (four teams of eight individuals) were given general maritime, and role specific training. Each team completed a high and low demand version of three scenario types: return to periscope depth, inshore operations, and dived tracking. On completing each scenario, the workload of participants was assessed using an electronic version of the Bedford scale. Preliminary results suggest the workload of operators was affected by scenario demand and type. Results also suggest that the co-location of operators had a positive effect on the demand placed on them. The results are discussed, along with future analysis plans.

KEYWORDS

Submarine, workload, command team

Introduction

Submarine control rooms represent a high state of evolution, from decades of operation, however this does not mean that they cannot be improved (Stanton, 2014). Larger volumes of data are being generated by more advanced sensors because of technological advancements (Dominguez, Long, Miller & Wiggins, 2006). Submarine command teams of the future will need to manage this data and maintain suitable levels of workload to achieve optimal performance (Roberts, Stanton & Fay, 2017).

Distributed cognition considers cognition beyond the scope of the individual, to include interactions between people and resources in the environment (Hollan, Hutchins & Kirsh, 2000). Cognitive processes may be dispersed among members of a team and require coordination to accomplish tasks (Hollan et al., 2000). This approach can be used when considering a socio-technical system; where social and technical factors interact in goal-directed behaviours (Walker, Stanton, Salmon & Jenkins, 2007). A submarine control room is an example of a socio-technical system with distributed cognition, as information must be gathered from multiple operators and sensors, to form an overview of a tactical picture (Stanton, 2014).

Levels of workload are a key consideration for safety and efficiency in a socio-technical system (Gregoriades & Sutcliffe, 2008). Workload is defined as demand placed on an individual that uses finite cognitive resources such as attention and processing capacity (Gregoriades & Sutcliffe, 2008). Overload, or under stimulation of workload could lead to operator error or performance decline, and ultimately, performance failure (Ayaz et al., 2012). Consequently, the assessment of workload in a submarine control room is needed to ensure safe operation and to prevent performance failures causing accidents. The Bedford scale, which considers levels of workload in terms of 'spare capacity', can be used to provide subjective ratings of workload (Roscoe & Ellis, 1990). Roscoe and Ellis (1990) found that it was possible to distinguish between different flight tasks using the Bedford scale, and concluded that it had good applications in practical settings.

Method

The first phase of the Command Team Experimental Testbed (ComTET) programme was to a build a mid-fidelity submarine control room simulator (Roberts, Stanton & Fay, 2015). The second phase of the programme was to conduct experiments within the simulator (Roberts et al., 2017). Stanton and Roberts (2017) conducted baseline testing within the simulator to establish current functionalities of a submarine command team. The recommendation to co-locate the Sonar Controller (SOC) and the Operations Officer (OPSO) to alleviate load between them formed the basis for Manipulation One (co-location). A preliminary assessment and comparison of subjective workload scores from baseline and the co-location condition are presented here.

Figure 1: The ComTET submarine control room simulator

A detailed description of the ComTET (see figure 1) build process and capabilities is given by Roberts et al., (2015), but a brief overview is given here. The ComTET simulator consists of nine networked stations, at which participants can play one of several roles; an OPSO, SOC, Sonar Operator (x2), Target Motion Analysis Operator (x2), a Periscope Operator, and Ship Control Operator. A member of the ComTET team played the final role, Officer of the Watch (OOW). Six scenarios (see table 1) were designed for use in the Dangerous Waters simulation engine, and all teams completed all scenarios. Subject Matter Experts (SMEs) were involved in designing all scenarios, to capture a representative range of submarine operations. The order of scenario completion was counterbalanced across the 4 teams, with scenario length averaging 45 minutes.

The ComTET team also developed a tutorial package to train participants in their randomly assigned roles. Tutorials included an overview of basic concepts such as 'bearing', 'course', 'speed', and 'range,' as well as operator specific tutorials (e.g., sonar tutorial). A communications game was also designed, where participants practiced military verbal protocol. This taught participants how to organise communications with other team members, such as how to initiate verbal communication with another operator.

Table 1: Description of scenarios

	Demand	*Number of contacts*	*Description*
Return to Periscope Depth (RTPD)	Low	4 – Fishing	RTPD to transmit intelligence within a large temporal window. All contacts (objects detectable by the submarine) need to be located and ranged to find an optimal route to RTPD.
	High	9 – Fishing 3 – Catamaran 1 – Biological (whale)	Severe submarine malfunction so must RTPD as quickly as possible. All contacts ranged to find best route to RTPD.
Inshore Operations (INSO)	Low	3 – Merchant 1 – Yacht 1 – Freighter	Gather intelligence on a building after navigating inshore. Scenario ends once periscope can photograph the building on land.
	High	2 – Merchant 1 – Powerboat 5 – Fishing	Identify and track a 'suspicious' vessel inshore to gather intelligence.
Dive Tracking (DT)	Low	3 – Fishing 1 – Sailboat 1 – Nimitz	Locate and track the priority contact, the Nimitz (warship). Scenario is complete when the Nimitz has been tracked and all other contacts have solutions.
	High	7 – Fishing 2 – Merchant 1 – Nimitz	Priority contact needs to be located and tracked after emergency go deep procedure.

ComTET was equipped to collect data from various sources and perspectives, and a mix of objective and subjective data was collected (for a full list see Roberts et al., 2017). The Bedford scale was administered electronically to assess subjective levels of workload (Roscoe & Ellis, 1990). This was a scale from one to ten, with one indicating low workload, and ten indicating high workload. Participants could enter decimals or whole numbers to indicate their level of workload.

Participants attended the ComTET facility for a two-day period; one training day and one testing day. Participants were recruited opportunistically using posters and by contacting local groups with a relevant interest. Four teams of eight individuals (student and industry teams) completed baseline testing, with a further four teams recruited for the co-location condition. Previous research has concluded that the use of novice teams is justified, with few significant differences between them (Walker, Stanton, Salmon, Jenkins, Rafferty & Ladva, 2010). The study protocol received ethical approval from the Ministry of Defence Research Ethics Council (MODREC) (Protocol No: 551/MODREC/14) and the University of Southampton Research Ethics Committee (Protocol No: 10099).

On the first (training) day, informed consent was obtained from the participants before commencing with the training tutorials and communications games. In the afternoon, participants took part in practice scenarios. Participants were given regular breaks and provided with sustenance on both days.

On the second (testing) day, participants were required to complete a final practice scenario. This allowed experimenters to assess whether adequate performance was being maintained, and for the provision of any additional training. Here, adequate performance refers to whether participants were able to perform the tasks required for their role, for example a sonar operator being able to detect and designate contacts. All scenarios were completed, with breaks at suitable intervals. Each scenario started with the OOW providing a scenario briefing, detailing overall objectives. The

OOW then tactically guided each scenario, as would occur in a submarine control room. After each scenario, participants were presented with an electronic version of the Bedford scale to assess workload (Roscoe & Ellis, 1990).

Results

The means and standard deviations of baseline Bedford scores for high and low demand scenarios are presented in table 2 and figure 2. The RTPD scenarios had the largest difference between mean Bedford scores, with scores greater in the high demand (*M* = 6.47, *SD* = 2.57) than in the low demand (*M* = 3.41, *SD* = 1.27). When comparing all scenarios, the RTPD low demand also had the lowest score, whilst the RTPD high demand had the highest.

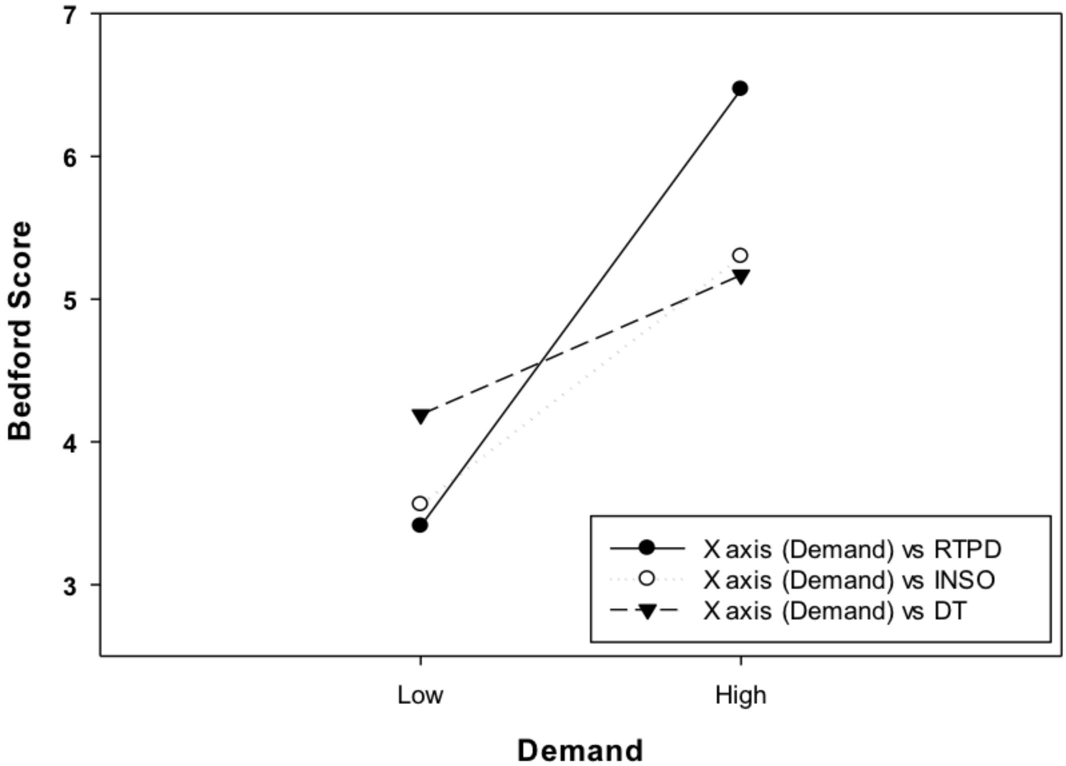

Figure 2: Baseline Bedford scores for RTPD, DT, and INSO low and high demand

In the INSO scenarios, the Bedford scores were greater in the high demand (*M* = 5.30, *SD* = 2.68) than in the low demand scenario (*M* = 3.56, *SD* = 1.99). Similarly, in the DT scenarios the low demand Bedford scores (*M* = 4.19, *SD* = 1.64) were lower than the high demand (*M* = 5.17, *SD* = 2.33). The mean scores for the DT scenarios high and low were the closest out of the three scenarios.

Table 2: Means and SDs of Baseline Bedford scores for high and low demand scenarios

	Low demand Mean ± SD	*High demand Mean ± SD*
RTPD	3.41 ± 1.27	6.47 ± 2.57
INSO	3.56 ± 1.99	5.30 ± 2.68
DT	4.19 ± 1.64	5.17 ± 2.33

The means and standard deviations of co-location (Manipulation One) Bedford scores for high and low demand scenarios are presented in table 3 and figure 3. Like baseline, the RTPD scenarios had the largest difference between mean Bedford scores, with scores greater in the high demand (*M* = 5.78, *SD* = 1.90) than in the low demand (*M* = 2.65, *SD* = 1.15). Additionally, the RTPD low demand had the lowest score of all the scenarios, and the RTPD high demand had the highest.

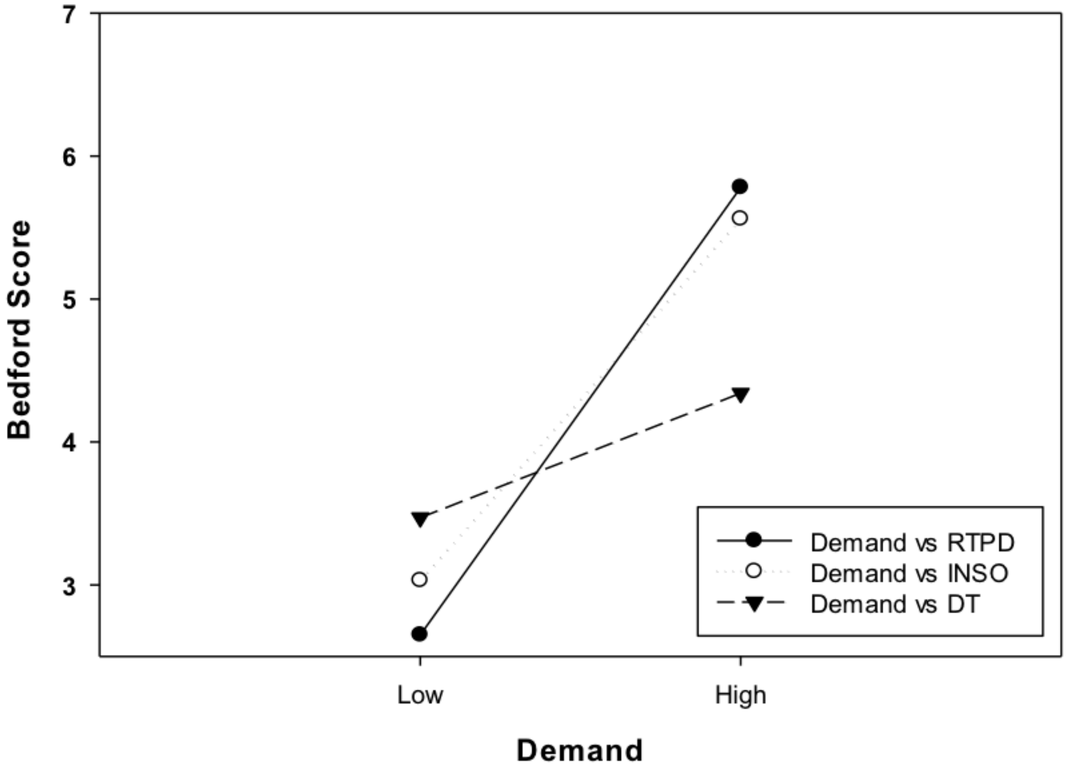

Figure 3: Co-location Bedford scores for RTPD, INSO, and DT low and high demand

In the INSO scenarios, the Bedford scores were greater in the high demand (*M* = 5.56, *SD* = 1.85) than in the low demand scenario (*M* = 3.03, *SD* = 1.40). In the DT scenarios, the low demand Bedford scores (*M* = 3.47, *SD* = 1.48) were lower than the high demand (*M* = 4.34, *SD* = 1.68). Consistent with baseline, the scores for the DT scenarios high and low were the closest out of the three scenarios. The Bedford scores for the high demand INSO are the only scores that increased in the co-location scenarios, all other Bedford scores decreased in the co-location condition compared to baseline.

Table 3: Means and SDs of Co-location Bedford scores for high and low demand scenarios

	Low demand Mean ± SD	*High demand Mean ± SD*
RTPD	2.65 ± 1.15	5.78 ± 1.90
INSO	3.03 ± 1.40	5.56 ± 1.85
DT	3.47 ± 1.48	4.34 ± 1.68

Discussion

The aim of the current work was to examine workload subjectively with a submarine command team. Differences in baseline Bedford scores between high and low demand scenario types indicate scenarios were adequately designed to elicit differing workload levels. Also, differences between scenario types suggest that the experimental design successfully captured a range of operational procedures and objectives. This aligns with previous analysis conducted by Roberts et al. (2017) of subjective (e.g., NASA-TLX) and objective (e.g., temperature) data. The co-location configuration (Manipulation One) had a clear impact on the Bedford scores of the participants. In all scenario types (except INSO high), the scores reduced which may be indicative of a reduced load on the participants. The increase in workload scores in the INSO high demand may be due to high reliance on the periscope operator in this scenario, whose position in the command room was not changed between Baseline and Manipulation One. Thus the wait for information from this operator was still present. Despite the reduction seen in workload between Baseline and Manipulation One, differences still exist between high and low demand, indicating that the experimental design holds for Manipulation One. Whilst the current work did not test expert teams, previous research comparing novice and expert teams found no significant difference in workload scores (Walker et al., 2010). So it may be possible to generalise the results found in the current study.

The current work provides subjective verification for the experimental design. In both Baseline and Manipulation One, differences between high and low demand were observed. Furthermore, the co-location configuration appears to have had a positive effect on the reduction of Bedford scores. This can be used as subjective verification that the new configuration reduces the load placed on the operator. The co-location configuration may have facilitated collaboration between operators better than baseline, so affecting participant's perceptions of workload. Whilst there are clear differences in Bedford scores between baseline and the co-location configuration, a greater number of teams would allow for empirical comparisons, which would provide evidence for future changes. The ComTET team intends to collect further data to allow for significance testing.

References

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage, 59*(1), 36-47

Dominguez, C., Long, W. G., Miller, T. E., & Wiggins, S. L. (2006, June). Design directions for support of submarine commanding officer decision making. In *Proceedings of 2006 Undersea HSI Symposium: Research, Acquisition and the Warrior* (pp. 6-8)

Gregoriades, A., & Sutcliffe, A. (2008). Workload prediction for improved design and reliability of complex systems. *Reliability Engineering & System Safety, 93*(4), 530-549

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI), 7*(2), 174-196

Roberts, A., Stanton, N., & Fay, D. (2015). The command team experimental test-bed stage 1: design and build of a submarine command room simulator. *Procedia Manufacturing*, *3*, 2800-2807. doi: 10.1016/j.promfg.2015.07.745

Roberts, A., Stanton, N. A., & Fay, D. (2017). The Command Team Experimental Test-Bed Phase Two: Assessing Cognitive Load and Situation Awareness in a Submarine Control Room. In *Advances in Human Aspects of Transportation* (pp. 427-437). Springer International Publishing. doi: 10.1007/978-3-319-41682-3_36

Roscoe, A. H., & Ellis, G. A. (1990). *A subjective rating scale for assessing pilot workload in flight: A decade of practical use* (No. RAE-TR-90019). Royal Aerospace Establishment Farnborough (United Kingdom)

Stanton, N. (2014). Representing distributed cognition in complex systems: how a submarine returns to periscope depth. *Ergonomics, 57*(3). 403-418. doi: 10.1080/00140139.2013.772244

Stanton, N. A., & Roberts, A. P. J. (2017). Examining Social, Information, and Task Networks in Submarine Command and Control. *IEEE Transactions on Human-Machine Systems*

Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science*, *9*(6), 479-499. doi: 10.1080/14639220701635470

Walker, G. H., Stanton, N.A., Salmon, P.M., Jenkins, D. P., Rafferty, L., & Ladva, D. (2010). Same or different? Generalising from novices to experts in military command and control studies. *International Journal of Industrial Ergonomics, 40*(5), 473-483. doi: 10.1016/j.ergon.2010.04.003