VICTORIA Class Submarine Human-in-the-Loop Experimentation Plan

A. Hunter, M. Hazen, and T. Randall,
Defence Research Development Canada Atlantic Research Centre

Abstract

The DRDC Atlantic Research Centre has designed an Integrated Information Display (IID) to aid the warfighting capabilities of the Commanding Officer (CO) and Watch Leader (WL) aboard the Canadian Forces (CF) VICTORIA Class Submarines (VCS). Having completed a usability assessment on the IID, the next step was to assess the impact of the IID on warfighting capabilities during human-in-the-loop (HIL) experimentation. It was hypothesized that the IID would increase warfighting capabilities and facilitate accurate and timely decision making while on watch. HIL experimentation took place in the VICTORIA Class Experimentation Laboratory (VCEL) which is a built-to-scale mock-up of the VCS operations room (ops room). The VCEL facility emulates an actual VCS ops room with real consoles, chairs, lighting, computer displays, layout constraints, and periscopes. During experimentation, the simulation environment, including sonar, Target Motion Analysis (TMA), helm, Electronic Support Measures (ESM), fire control and periscope displays were emulated by the serious game Dangerous Waters (DW). DW was used for scenario generation and it was used to feed simulated data to each of the displays, including the Flash-based IID which was developed at DRDC. Scenarios for the HIL experiment, developed by subject matter experts (SMEs), required the CO/WL to position the submarine in an optimal location to build the Recognized Maritime Picture (RMP) and identify contacts of interest (COIs). Considerable effort was made to operationally define measurable dependent variables with theoretical and practical implications specific to submarine warfighting capabilities. The experiment was also designed to effectively balance experimental control and ecological validity. In order to triangulate the results, our experimental plan involved integrating various qualitative and quantitative data sources, including eye tracking measurements, SME evaluations, questionnaires, personnel movement, and communication. This paper will outline the experimental plan, methodologies, VCEL, and the challenges involved in designing complex HIL experiments.

Introduction

Command team personnel aboard submarines are faced with unique Command and Control (C2) challenges imposed by the limitations of crew size and operating environment. Submarines are often described as operating "blind" with a high level of uncertainty which puts many constraints on the decision making capabilities of the command team (Hautamaki, Bagnall, & Small, 2005; Dominguez, Long, Miller, Wiggins 2006; Kirschenbaum & Arruda, 1994). Research on submarine command teams suggests a number of difficulties related to data uncertainty and assimilation, environmental uncertainty, as well as team related issues such as workload and communication (Hautamaki et al, 2005; Jones, Steed, Diedrich, Armbruster & Jackson, 2011). To better understand the specific challenges of the Canadian Forces VICTORIA Class submarines (VCS) command team, Defence Research Development Canada (DRDC) Atlantic Research Centre completed a number of Cognitive Work Analyses (CWA) - including a Work Organization Analysis, a Cognitive Transformations Analysis (CogTA) and a Strategies Analysis (Chalmers, 2010, 2011). These analyses revealed a number of C2 challenges related to the ease with which the Commanding Officer (CO) and Watch Leader (WL) were able to assimilate mission-relevant information to aid effective warfighting performance. These findings are in line with the cognitive difficulties identified by Dominguez et al (2006) who also found that the CO had difficulties assimilating information and managing the uncertainty inherent in submarine operations.

To address these challenges, DRDC Atlantic Research Centre identified the need for an Integrated Information Display (IID) to improve the accessibility and integration of information relevant to tactical command. The IID was designed for use by the CO and WL under the hypothesis that by reducing the amount of cognitive effort required to assimilate and integrate information from multiple systems spread around the control room, the officers would have more cognitive resources available for tactical command reasoning and planning. Improved tactical reasoning was expected to result in improved warfighting performance. Contracted human factors engineering resources were used extensively in the development of the IID and the resulting reports are under review for releaseability. The IID was designed by a team of six Human Factors (HF) analysts (Bruyn-Martin, Taylor, & Karthaus, 2009). The design team used a modified form of Ecological Interface Design (EID) that utilized information from the CWA to develop content and inform the spatial layout of the display (Chalmers, 2011). The intent of EID is to support the demands of the human instead of just presenting data. Following HF design standards (MIL-STD 1472G) the team developed new graphical representations that integrated information currently available at multiple locations within the control room. All iterations of the design were evaluated by two submariner SMEs. Integrating elements in a single display was expected to reduce the perceptual and attentional requirements leading to a decrease in workload and increase in the efficiency of decision-making in the control room.

After a functional prototype was complete, the display underwent usability and HF standards assessments (Hunter & Randall, 2013; Lamoureux, Pasma, & Kersten, 2012). Feedback from the assessments was prioritized and changes to the IID display were made. The updated experimental version of the IID was dynamic and interactive with the intent of providing a "functional picture" to support the work demands of the CO and WL (Chalmers, 2011). It was hypothesized that the IID would improve the warfighting capabilities of the CO and the WL in such a way that it would facilitate accurate and timely decision-making. Following the usability analysis and SME validation a number of changes in organization were made prior to experimentation. The final version of the display is shown in Figure 1.

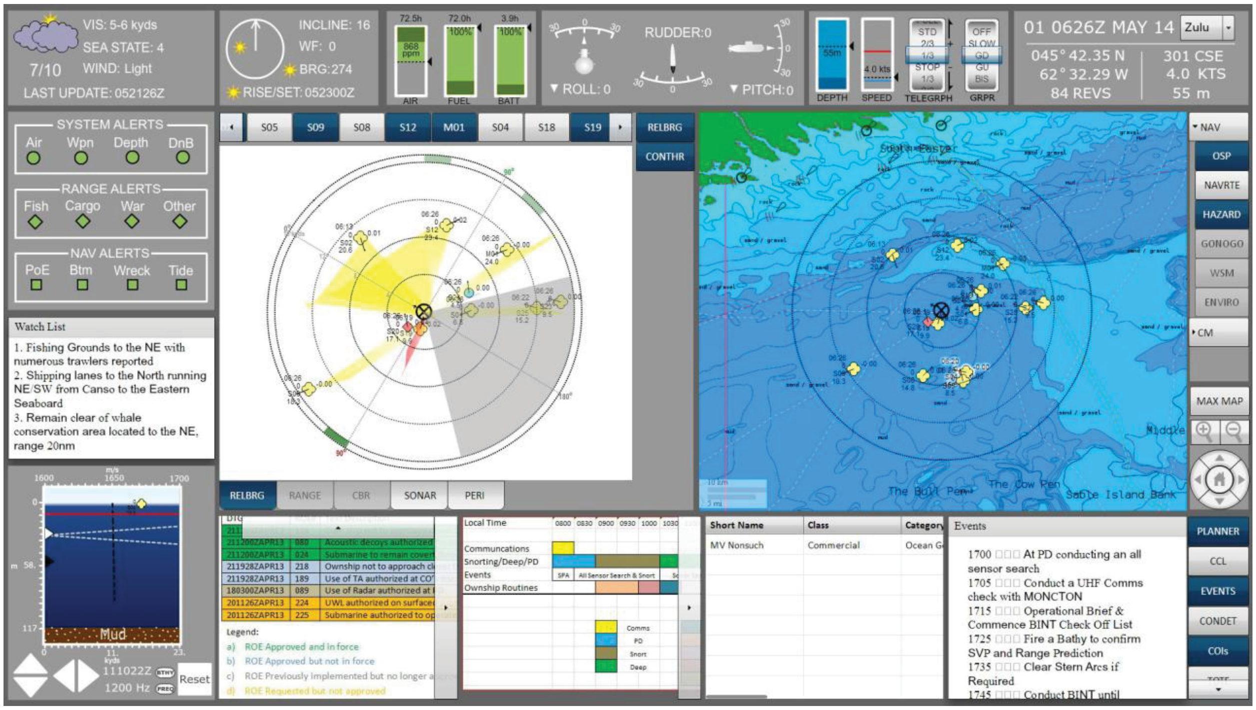


Figure 1. Final version of IID used in experimentation.

Experimental Environment

The next step in this project was to evaluate the IID during human-in-the-loop (HIL) experimentation where the IID, along with sonar, Target Motion Analysis (TMA), helm, Electronic Support Measures (ESM), fire control, periscope and the ECPINS display, were fed by the serious game Dangerous Waters. All of this took place in an experimental facility known as the VICTORIA Class Evaluation Laboratory (VCEL) (see Figure 2) (Hazen, Gillis, Coady, Franck, & Dillman, 2014). Given the unique space and layout restrictions in the VCS, and the likely impact these restrictions have on human behavior, VCEL was designed to be as representative of the actual VCS control room as possible. Details such as overhead bars, chairs, cabinets, red lighting, computer displays, distance between operator stations, and the placement and size of the periscopes were all taken into consideration.

Figure 2. Exterior view of VCEL.

Three locations for the placement of the IID in VCEL were considered- the chart table, above the CO's chair and between sonar and fire control. Using a cardboard cut-out of the IID the visibility and readability of the display was assessed at each of these three locations. While the CO has a dedicated location to sit in the control room, the WL does not. Based on our observations during training sessions the WL spent the majority of his time behind the fire control operators and at the periscope when the submarine was coming to periscope depth (PD). Based on all of this information it was decided that the best location for the IID was between sonar and fire control (see Figure 3).

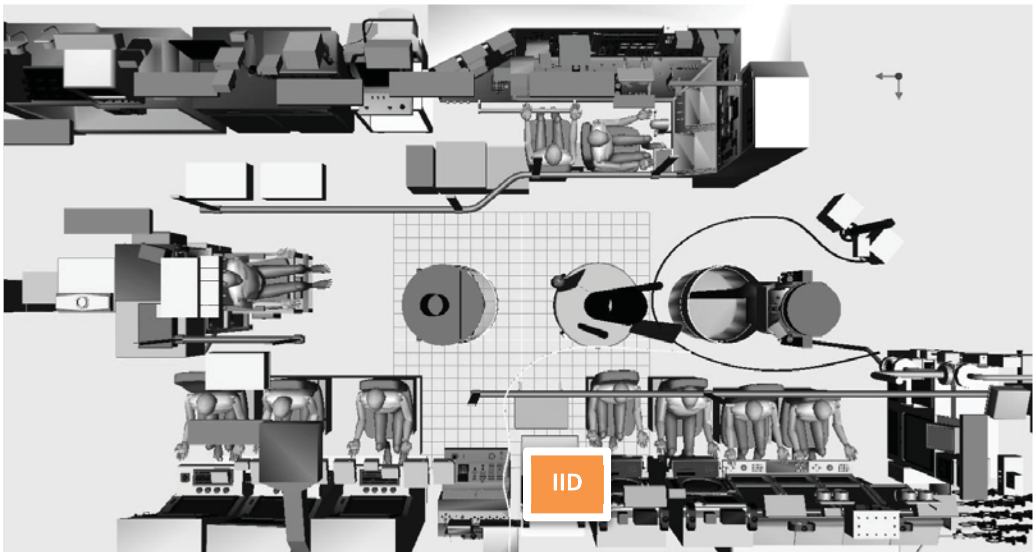


Figure 3. The placement of the IID in the VCS control room between sonar and fire control.

Experimental Design

The experiment was conducted over the course of three days in March/2014. Day 1 was used to obtain consent, collect demographics information, and train all members on the appropriate simulated operator stations (sonar, fire control, navigation) and the IID. Day 2 was used to run through the control and experimental condition twice with Team 1, and similarly Day 3 repeated the runs of Day 2 with Team 2. Each simulation run lasted for approximately 1.5 hours, and was followed by the completion of a debriefing questionnaire for the team members and a separate one for the WL. At the end of the 4 runs there was a 20 minute debriefing session to elicit feedback from the team members and the WL.

While tests of statistical significance will be limited by the small sample size it is still worth reviewing the design that was used. The independent variable of interest was condition. A repeated measures design was employed. Every effort was made to control extraneous team and scenario variables (described below). In order to test for the effect of these variables on behaviour scenario and team were treated as independent variables for analysis purposes.

		Within Subjects Variable			
Between Subjects Variable	*Team 1*	Control Condition [1]	Experimental Condition [2]	Experimental Condition [3]	Control Condition [4]
	Team 2	Experimental Condition [4]	Control Condition [3]	Control Condition [2]	Experimental Condition [1]

[1]Scenario 1
[2]Scenario 2
[3]Scenario 1a
[4]Scenario 2a

Table 1. Experimental design.

Team. While the focus of the experiment was on the WL was important that the appropriate control team members were present in the VCEL control room to support the flow of information. For this experiment we requested two separate teams with similar levels of experience to minimize the chances of a significant team effect.

The average age of Team 1 was 35.7 years and Team 2 was 33.8 years. Team 1 had on average 7.5 years of submarine experience, while Team 2 had an average of 6 years of submarine experience. Team 1 had an average of 1.9 years of operational experience working with the team they were assigned to in the experiment and Team 2 had an average of 1.2 years. The WL on Team 1 had 4.5years of experience in this position, whereas the Team 2 WL had only 3 months experience. Initial analyses suggest that the difference in experience was not an issue but this will be taken into consideration when making comparisons between the two teams. It should also be noted that due to the limited availability of submariners the ESM and sonar operator participated as part of Team 1 and Team 2. No carryover effects are expected due to the noncritical roles of these two participants. Due to limited availability we were also unable to get a CO.

Team Member Selection. To ensure that we solicited the appropriate team members for participation we used an information flow analysis. The information flow analysis, as can be seen in Figure 4, was used to identify key players in the passing and receiving of information with respect to the CO and WL.

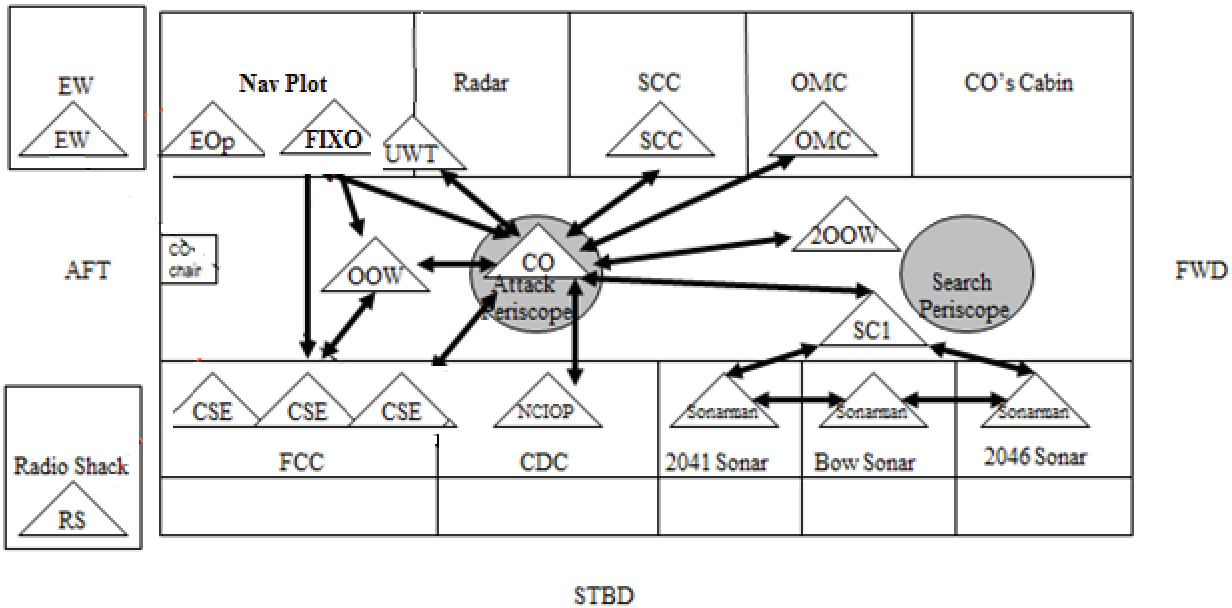


Figure 4. Information flow analysis in the VCS control room during an intelligence, surveillance and reconnaissance mission. Arrows indicate the flow of information. Adapted from Taylor, Karthaus, & Bruyn-Martin (2009).

Based on this information, the participant recruitment list included a fire control operator (Combat Systems Engineer (CSE)), a naval combat information operator (NCIOP) stationed at the command display console (CDC), a senior sonar supervisor (SC1) stationed behind the 2040, a sonar operator stationed at the high frequency (HF) 2040 sonar console, a helmsman (OMC) stationed at the helm, a Navigation Officer (NavO which is equivalent to the FIXO in the above diagram), a WL (aka Officer of the Watch OOW) and a CO. The underwater telephone (UWT) is typically operated by the Sonar operator that is responsible for the low frequency (LF) radar. In this experiment, the UWT was not functional and as a result this position was scripted by an SME as the scenario required it. The system control console was not manned as VCEL does not support consoles required for the performance of that particular job.

Condition (Within Subjects Variable). In both the control and experimental condition, the operators were asked to carry out their mission in the way that they normally would while on a VCS. The systems the operators were using were slightly different than those on the VCS but the tasks were essentially the same. The operators were required to process data (e.g., initiate sonar tracks, perform target motion analysis) and provide information to the WL verbally (as they would during real operations). During the experimental runs, the WL was asked to use the IID to satisfy his information requirements. In the control conditions, the WL was asked to carry out his mission in the way that he normally would while on a VCS.

Scenario

In this experiment we have two main scenarios and two scenarios that are modifications of the main scenarios (1a and 2a). While the original plan was to do a repeated measures using the same scenarios we consulted with the SMEs and they suggested adding two additional scenarios so that each run was unique. The two additional scenarios were simple variations on scenario one and two where the location of contacts changed but the submarine patrol area and number of contacts remained the same. All four the scenarios were designed by two subject matter experts (SMEs)-one a retired Commanding Officer and the other a retired OBERON class Sonar Operator. All of the scenarios required the submarine to perform a surveillance mission. At a tactical-level this requires the CO/WL to place the submarine in an optimal position to build the Recognized Maritime Picture (RMP) and identify contacts of interest (COIs). Scenario one, known as the "oiler" scenario, required ownship to conduct a BINT and perform a covert surveillance operation of an at sea fuel transfer between an oil tanker and a submarine. In scenario two, known as the "smuggling" scenario, ownship was required to coordinate a covert surveillance and intelligence gathering mission on a trawler suspected to be smuggling members of a terrorist group. Ownship was responsible for reporting and delivering all data to a Maritime Coastal Defence Vessel (MCDV) with onboard law enforcement personnel.

Scenario Development

Scenario development was one of the most challenging and time-consuming aspects of experimentation planning. In order for us to minimize the likelihood of a significant scenario effect, we had to ensure that the two main scenarios (1 and 2) were similar in their complexity, difficulty, workload and workflow. To facilitate this, both scenarios had the same number of contacts, sensors, mission type, mission goals, BATHY information, sea state, weather, time of day, standard operating procedures (SOPs), and Intel. In both scenarios, everything except ownship motion was scripted. To determine if these details were effective in making the scenario equivalent in communication a communication analysis was completed (Kersten-Kwan, Bruyn Martin & Matthews, 2013).The communication analysis required two SMEs to role-play the various members of the command team (CO, WL, NavO, SonSup, ESM, FC, Helm and Scanner Operator (ScanOp)). The scenarios were run on a large computer screen in real-time during which the SMEs followed communication protocol and verbalized communication as it would happen with respect to each position. During these sessions the frequency, type, context and semantic content of communications were recorded on a time stamped spreadsheet. Unfortunately we did not have time to perform the same assessment on scenario 1a and 2a, but given that these scenarios were modifications of the main scenario it is expected that they would not differ. The data will be evaluated for a scenario effect by treating the extraneous scenario variable as an independent variable.

Frequency of Communication. The results of the communication analysis indicate that communications by command position between the two scenarios share a similar trend but are not identical. The results suggest that the CO, WL, Sonar Supervisor and ESM operator have the greatest frequency of communication. While the communication frequency trend is similar, there are notably more communications in the smuggling scenario. It should be noted that communication from the Helm was accidentally omitted in the Oiler scenario but it is expected that it would yield similar numbers as the smuggling scenario.

Communication Type. The next analysis involved an assessment of type of communication (adapted from Swain and Mills, 2003)- command, request, reply, recommend, report, information and brief. The distribution of communication frequency across these types of communication provides an indication of workflow. The analysis revealed that reports were the most common type of communication.

Scenario	Communication Types						
	Command	Request	Reply	Recommend	Information	Report	Briefs
Oiler	44	1	1	0	0	39	8
Smuggling	61	1	1	2	0	69	20
Total	**105**	**2**	**2**	**2**	**0**	**108**	**28**

Table 2. Frequency of communication type broken down by scenario (from Kersten-Kwan, Bruyn Martin & Matthews, 2013).

Communication Context. Next a comparison of communication context was completed. This evaluation also provided an indication of workflow with respect to the importance of information being passed around the control room. The following categories of communication content were used – ownship; contacts; environment; logistics; other. The bulk of the communications were about contacts while the remaining communications were related to ownship. This trend was consistent in both scenarios.

Scenario	Communication Context				
	Ownship	Contacts	Environment	Logistics	Unsure
Oiler	33	60	0	0	0
Smuggling	55	96	0	0	3
Total	**88**	**156**	**0**	**0**	**3**

Table 3. Frequency of communication context broken down by scenario (from Kersten-Kwan, Bruyn Martin & Matthews, 2013).

Communication Content x Position. When content was broken down in relation to position the results show that the CO in the oiler and smuggler scenario had nearly an equal split in communication that referenced ownship vs. contacts.

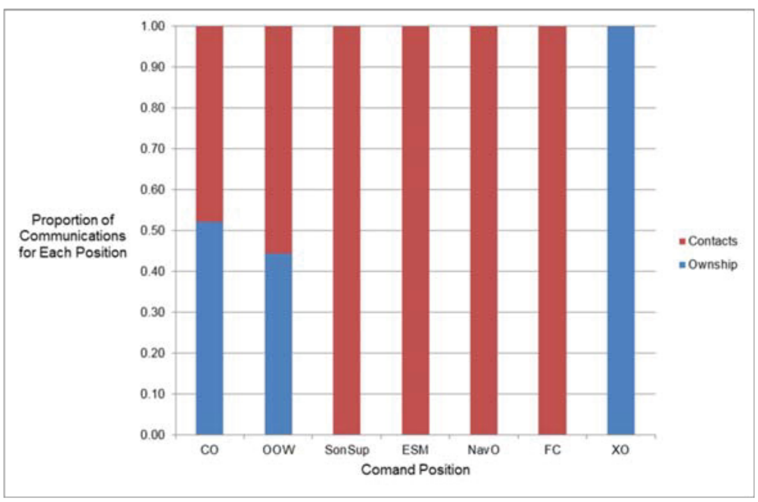


Figure 5a. Proportion of communication broken down by position and content for the oiler scenario (from Kersten-Kwan, Bruyn Martin & Matthews, 2013).

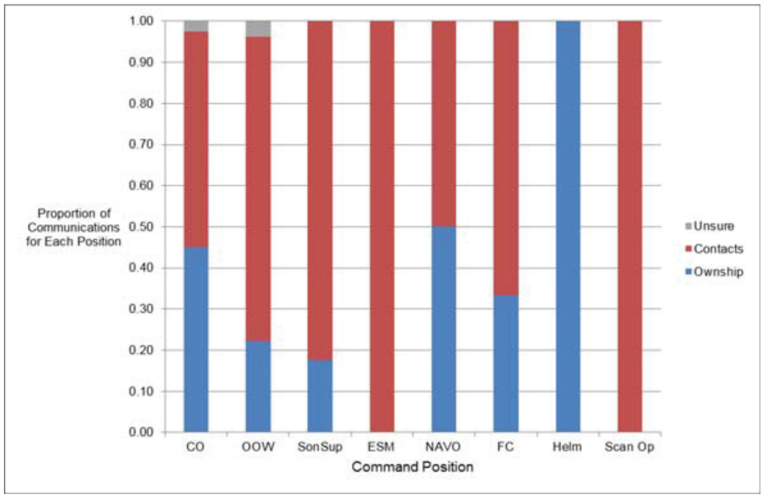
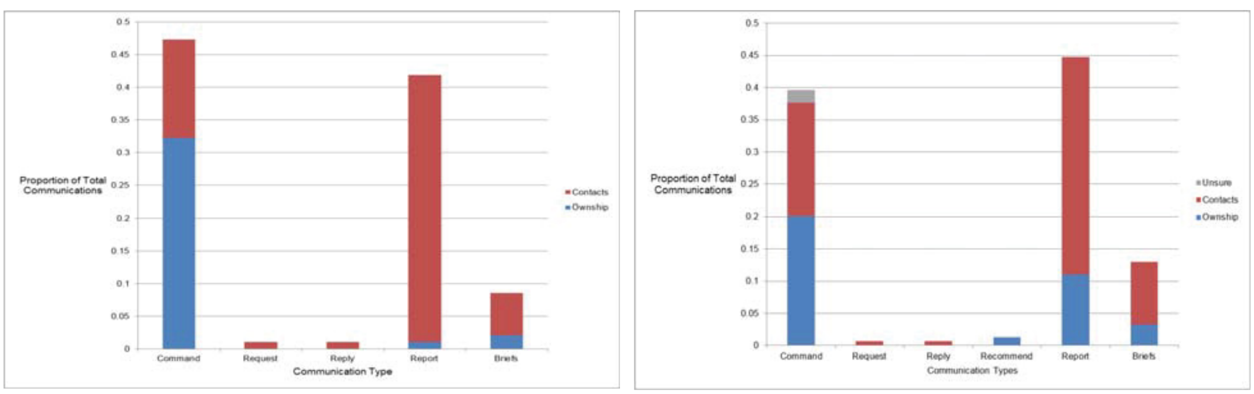


Figure 5b. Proportion of communication broken down by position and content for the smuggler scenario (from Kersten-Kwan, Bruyn Martin & Matthews, 2013).

Communication Type x Content. As can be seen in the graphs below, the two scenarios follow similar trends when communication type is broken down by communication content. The majority of the reports are in reference to contacts, whereas commands are split between information pertaining to ownship and contacts.



a. b.

Figure 6. Proportion of communications by type and context collapsed over command position for (a.) “smuggling” scenario and (b.) “oiler” scenario (from Kersten-Kwan, Bruyn Martin & Matthews, 2013).

Scenario Comparison. With the exception of the smuggler scenario having more overall communication, the communication trends with respect to position, content, context, and type were similar between the scenarios. In the end we were satisfied that the scenarios were similar enough in workflow and communication for experimentation purposes.

Data Recording

For data collection purposes, VCEL was equipped with four wall-mounted video cameras. Each camera was used to record video from different angles throughout each 1.5 hour run. Videos were displayed 'live' on a large display outside of VCEL for viewing by the research team, including two SMEs, in order to guide/refine post-run questioning and obtain expert opinion on mission performance, and were stored to local hard drives for post hoc analyses. In addition, all participants were asked to attach a small MP3 to their shirt. A distinctive audio tone was used to signify the start of each run and provide a way to coordinate the recordings of one operator with another in the post-study analyses. Also four Microsoft Kinect sensors were set-up in VCEL to measure movement in and around the control room. The WL was also asked to wear a pair of lightweight (100 grams) SensoMotoric Instruments (SMI) eye tracking glasses. Due to technical issues with the mobile recording device a laptop in a backpack had to be used instead. The glasses required 1-2 minutes of calibration with the WL prior to each run. Once all of the data recording equipment was turned on, the WL and his team were told to begin the experiment. They were the only ones in the VCEL control room during the experiment.

Dependent Variables

As mentioned above, our focus was on measuring the change in warfighting effectiveness between the control and experimental condition. Finding objective measures of warfighting performance was difficult since the sparse literature on performance metrics is often specific to the systems and scenarios used in those particular cases and, given the exclusivity of the submariner population, there is little publicly available submarine related literature to draw from. To aid us in developing a list of dependent variables, a submariner SME session was held to assess how decisions are made and what elements are used in decision making (Lamoureux, Pasma & Kersten, 2011). This section will outline the relevant dependent variables gathered from the referenced report and the open literature.

Scenario Based Metrics. The following scenario relevant metrics were derived from the SME session (Lamoureux, Pasma & Kersten, 2011) and the open literature. The metrics will be calculated by the experimentation team during the analysis phase using the recorded ground truth data for comparison. It is hypothesized that the IID will aid improved performance (accuracy and efficiency) on each one of the following metrics. The metrics are broken down by the element of submarine mission that they support.

1.) Contact management metrics- Number of lost contact incidences, number of contacts detected vs. number in scenario, number of contact re-classifications, false alarms, or repeated contacts, and the accuracy of target motion analysis (TMA) when compared to ground truth, range to target time (adapted from Kirschenbaum & Arruda, 1994), and solution completeness (Huf et al., 2004).

2.) Covertness metrics- Time spent at periscope depth (PD), number of counter detections, and frequency of cavitation.

3.) Planning metric- Duration of the mission vs. the planned mission.

4.) Safety metrics- Collisions with vessels or land, accuracy of closest point of approach (adapted from Hautamaki et al.,2005), look interval duration, frequency of going deep, and accuracy of pilotage.

SME evaluations. In addition to the objective measures presented above we also wanted to assess why particular decisions were made or were not made. This is difficult since domain experts often apply heuristics gained from experience to make decisions (Chi, 2006). In order to better understand why WL may have performed a certain action or made a particular decision we enlisted the help of two SMEs. During the experiment, the SMEs were located outside of VCEL with access to real-time video and audio streams. They were asked to take notes while paying particular attention to behaviours or communications that are worth following up after the experiment. In addition, the SMEs were asked to fill out a questionnaire at 30-minute intervals (see Table 4 below). This method is similar to one used by Jones et al (2011), wherein they had reviewers rate performance on a likert scale. The SMEs were asked to make their assessments separately to reduce bias and increase the reliability of responses.

SME Assessment Questionnaire

Please use the scale indicated below to rate the questions. Please base your answers on behaviours that occurred in the last 30 minutes.

	1 Poor	2 Fair	3 Average	4 Good	5 Excellent	N/A
WLs consideration of overall safety.						
WLs consideration of covertness.						
WLs overall decision-making efficiency.						
WLs timeliness of decision-making (did he respond to situations in a timely manner).						
WLs understanding of the tactical picture.						
WLs assignment of priority to contacts.						
WLs management of ownship signature.						
WLs achievement of mission milestones.						
WLs adherence to the plan.						
WLs effective and timely communication of the plan, or change in the plan, to the team.						
WLs situation awareness.						
WLs workload.						
Teams mutual understanding of the mission goals and how to achieve them.						
Team workload.						
WLs utilization of the IID (only for IID condition)						

Table 4. SME assessment questionnaire to be given at 30-minute intervals.

Communication. Another way to assess performance is by evaluating communication between team members. For this experiment we are interested in communication flow to and from the WL. It is hoped that changes in communication between the two conditions should provide an indication of the added benefits and/or gaps left by the IID. A communication flow analysis will be completed for the control and experimental conditions for each scenario. The data collected will be categorized according to communication type – command, request, reply, report, information sharing, and brief (adapted from Swain and Mills, 2003). The content will be categorized as being related to ownship, contacts, environment, logistics and miscellaneous. Sender and intended receiver of the communication will also be evaluated.

We hypothesize that in the experimental condition the WL will make fewer requests for information. This is expected as the WL will have access to more information in the IID condition in comparison to the control condition, which should reduce the need for requesting information. It is possible that the WL will issue more commands in the IID condition as he has greater access to information. Given the nature of the scenario, we expect the proportion of content to remain mostly unchanged with the majority of communication focused on contact related information.

WL Movement. Changes in the movement of the WL between conditions will be assessed. As mentioned above, the WL does not have a dedicated space in the control room so where he locates himself is an indication of what information and what team members are important to him. We would expect to find a change in movement with the addition of the IID since it is expected the WL will spend more time at the IID and therefore reduce the amount of moving around he needs to do.

Eye Tracking. Eye tracking is a measurement tool that allows for objective evaluations of visual and cognitive processing. Eye tracking research is based on the eye-mind hypothesis which states that we direct our eyes to elements in our environment that are of interest to us or reflective of what we are thinking about (Just & Carpenter, 1980). This works particularly well in research programs, such as human computer interaction (HCI), where participants are viewing a specific stimulus such as a display. As mentioned above, we asked the WL to wear a pair of light-weight SMI eye-tracking glasses, as pictured in Figure 7., during the experiment. The SMI eye-tracking glasses provided the WL with the freedom to wonder around the control room as he normally would.

Figure 7. SMI eye tracking glasses and mobile recording device.

The eye tracking data will be used to do a detailed investigation on the various areas of interest on the IID. Each of the areas in the IID (depicted in Figure 1) will be defined as an area of interest for the eye tracking analysis. It is our goal to determine from the eye tracking metrics what areas of the display are used the most and what the patterns of usage are. To do this the following metrics will be evaluated-

Fixations. A fixation occurs when the eye stops moving and remains stationary (Poole & Ball, 2005).

Dwell Time. Fixations that have a longer duration are indicative of cognitive processing or encoding of information (Just & Carpenter, 1980). The average of all fixation durations within the specified areas of interest will be calculated. Dwell time information will be used to infer what areas of the IID evoked the most interest.

Fixation Frequency. The number of fixations in each area of interest will be calculated. Higher frequencies are indicative of more interest in that information (Poole, Ball, & Phillips, 2004).

Proportion of Fixations by Area of Interest. The proportion of fixations will be calculated for each area of the IID. Again, areas with higher proportions of fixations provide an indication of interest.

Saccades. A saccade is defined as a movement of the eye from one fixation to another (Poole & Ball, 2005).

Saccade Frequency. The number of saccades will be calculated. Saccades are related to searching where more saccades reflect more searching. Eye movements that are driven by intent have fewer saccades because the eyes are directed towards a specific area of interest (Goldberg & Kotval, 1999 as cited in Poole & Ball, 2005).

Saccade Amplitude. The amplitude of the saccade from one fixation to the next will be calculated. The greater the amplitude the more intent the eye movement is thought to be. Greater amplitudes indicate more attention to that area (Goldberg & Kotval, 1999 as cited in Poole & Ball, 2005).

Scanpath. A scanpath represents a full sequence of fixations and saccades. Scanpath theory stipulates that visual information in a display or image is encoded in a pattern that reflects how the image was pieced together (Foulsham et al., 2012). Of particular interest are the scanpaths between different areas of interest on the IID so that we can better understand the information gathering strategies used by the OOW in the IID condition.

Pupillometry. Pupillometry is the measure of pupil diameter to provide an indication of workload. Pupil diameter has been shown to increase during complex and difficult tasks (Tooley & Demczuk, 2010; Laeng, Sirois, Gredeback, 2012). This will be one of eye tracking measures that will be used to directly compare between the experimental and control condition.

Heat Map. A heat map is a visual representation of fixation activity. While it does not provide quantitative information, it does provide a qualitative overview of activity. Heat maps will be used to guide quantitative analyses.

Conclusion

The main goal of this experiment was to evaluate how behaviour changes in response to the IID. We are specifically interested in behaviours that are reflective of improvements in warfighting capabilities and we expect that the IID condition resulted in improved warfighting capabilities. The challenge with large HIL experiments is the maintenance of ecological validity while controlling for confounds. To do this we carefully developed our independent variables to reduce potential confounds while still allowing the WL control over the submarine so that we could collect realistic behaviours. The design we employed was selected to minimize carryover, practice, scenario and team effects as much as possible. Ideally the change in behaviours exhibited between the control and experimental conditions are only due to the addition of the IID. Overall, the IID experimentation used a combination of good overall design with combinations of objective and subjective measures to compensate for the small sample size. While the results of the experiment are unlikely to have sufficient statistical power for definitive results, it included sufficient data collection on participant behaviour to provide the sponsor and overall project with the information required to assess and/or modify the IID concept for future fleet employment.

The experiment was an important step in the development of Canada's maritime C2 experimentation capabilities, with the introduction of the use of the VCEL facility and new experimentation technologies such as the Microsoft Kinect sensors and SMI eye tracking glasses.

References

Bruyn Martin, L., & Taylor, T. E. (2010). Validation and prioritization of C3 concepts for the VICTORIA Class Submarine: Final Report. Informal communication.

Bruyn Martin, L., Taylor, T., & Karthaus, C. (2009). Informal communication.

Chalmers, B. (2010). A Work-Centred Approach to Optimizing Human-System Integration in Command and Control for the VICTORIA Class Submarine, Proceedings of Undersea Human Systems Integration Symposium 2010, American Society of Naval Engineers, July 2010.

Chalmers, B. (2011), *Developing an Information Integration Display for Submarine Command and Control*, Proceedings of Human Systems Integration Symposium (HSIS) 2011, American Society of Naval Engineers, October 2011.

Chi, M. T. (2006). Two approaches to the study of experts' characteristics. In The Cambridge handbook of expertise and expert performance, 21-30.

Dominguez, C. O., Long, W. G., Miller, T. E., & Wiggins, S. L. (2006). Design directions for support of submarine commanding officer decision making, *Undersea HSI Symposium*. Mystic, CT.

Foulsham, T., Dewhurst, R., Nystrom, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach. *Journal of Eye Movement Research*, 3, 1-14.

Goldberg, H. J., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.

Hautamaki, Bagnall & Small (2005). Human Interface Evaluation Methods for Submarine Combat Systems. Retrieved from- http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.6661&rep=rep1&type=pdf .

Huf,S., Arulampalam, S., Mansell, T., Tynan, A., Brown, G., and Manning, R. (2004). Investigating the network enabled conventional submarine II: A summary of Australian simulation research. Presented at the 9th International Command and Control Research and Technology Symposium.

Jones, E., Steed, R,, Diedrich, F., Armbruster, R., & Jackson, C. (2011) Performance-Based Metrics for Evaluating Submarine Command Team Decision-Making, in Foundations of Augmented Cognition. Directing the Future of Adaptive Systems Lecture Notes in Computer Science. 6780, 308-317

Just, M. A. & Carpenter P. A. (1980). A theory of reading: on eye fixations to comprehension. Psychological Review, 87, 329–354.

Kersten-Kwan, C., Bruyn Martin, L., & Matthews, M. (2013). Informal communication.

Kirschenbaum, S. & Arruda, J. (1994). Effects of graphic and verbal probability information on command decision making. The Journal of the Human Factors and Ergonomics society, 36, 406-418.

Lai, G. & Lamoureux, T (2011). Development of measures of effectiveness and performance from cognitive work analysis products. DRDC Atlantic Contract Report CR 2011-282.

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry-A Window to the Preconscious? Perspectives on Psychological Science. 7, 18-27.

Lamoureux, T., Pasma, D., & Kersten, C. (2012). Informal communication.

Poole,A. & Ball, L. J. (2006). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In Ghaoui, Claude (Ed.). Encyclopedia of Human Computer Interaction. Idea Group

Poole, A., Ball, L. J., & Phillips, P. (2004). In search of salience: A response time and eye movement analysis of bookmark recognition. In S. Fincher, P. Markopolous, D. Moore, & R. Ruddle (Eds.), People and Computers XVIII-Design for Life: Proceedings of HCI 2004. London: Springer-Verlag Ltd.

Rehak, L., Karthaus, C., Lee, B., Matthews, M., & Taylor, T. (2011a). Informal communication.

Rehak, L., Karthaus, C., Lee, B., Matthews, M., & Taylor, T. (2011b). Informal communication.

Swain, K. & Mills, V. (2003) Implicit communication in novice and expert teams. DSTO TN-0474.

Taylor, T., Karthaus, C., & Bruyn Martin, L. (2009). Informal Communication.

Tooley, K. & Demczuk, V. (2010). The use of an eye tracker in work analysis. TTCP Defence Human Systems Symposium, May 17-19, 2012 Sydney, Australia.

US Department of Defense (2012). Department of Defense Design Criteria Standard: Human Engineering. MIL-STD-1472G.