Augmenting Human Cognition with a Digital Submarine Periscope Human-Machine Interface

Steph Michailovs¹, Jessica Irons², Zach Howard¹, Stephen Pond¹, Megan Schmitt², Matthew Stoker³, Troy Visser¹, Jason Bell¹, Gavin Pinniger¹, Madison Fitzgerald¹, Sam Huf², Owen Carter¹, & Shayne Loft¹

¹School of Psychological Science, The University of Western Australia, Perth

²Undersea Command and Control, Maritime Division, Defence Science and Technology Group, Perth, Western Australia

³Ergonomie, Sydney, Australia

This research was supported by a Research Grant (MyIP 9277) from the Defence Science and Technology Group, Australia awarded to Loft, Michailovs, Visser, Bell, & Pinniger. Address correspondence to Steph Michailovs, The University of Western Australia, WA 6009 Australia, email: steph.chen@uwa.edu.au, or Shayne Loft, The University of Western Australia, WA 6009 Australia, email: shayne.loft@uwa.edu.au.

The data that support the findings of this study are available from the corresponding author, [SM and SL], upon reasonable request.

Abstract

Advances in opto-electronics enables replacement of traditional periscopes which display only a portion of the horizon, with digital periscopes, which can potentially display a full 360° panoramic digital representation of the horizon. One advantage of an increased field of view (FOV) might be to reduce memory requirements and improve integration of visual information over time, thereby improving detection of vessels (contacts) at sea. Another potential advantage of moving to digital periscopes is that imagery can be paired with digital aids that can replace or simplify tasks traditionally performed manually, such as contact identification. This study compared a low-FOV visual concept to a high-FOV concept (Experiment 1), and then traditional analysis tools with digitized analysis tools (Experiment 2). The high-FOV concept and digitized tools provided a range of perceived workload, perceived usability, and performance benefits to the human operator, and we conclude that digitally augmenting the periscope can potentially support submariner tasks.

Keywords: visual search, human-machine interface, periscope, submarine, workload, usability.

Introduction

Rapid advances in technology enable replacement of current submarine periscopes with digital periscopes that allow for 360° imagery and integration with other sensor and tactical systems. With this new technology, the periscope watchkeeper, previously tethered to the conventional periscope via eyepieces and handles, could potentially observe the external environment in a variety of ways, including on a computer display. This provides new opportunities to consider the periscope interface from a human-centred design perspective. To provide a framework for these innovations, however, it is important to consider not just operator task requirements, but also the cognitive and perceptual constraints of operators, highlighted by research in fields such as psychology and neuroscience.

The operator of a conventional periscope can usually see only a portion of the horizon, typically about 32° on low magnification and about 8° at high magnification (Walker, 2000). A full search of the environment requires physical movement, rotating the restricted periscope-viewing window around in order to scan the 360° environment. From the perspective of safeguarding the submarine (Ownship), the need to physically slew the periscope may delay the detection, classification, and localisation of contacts (i.e., vessels). From the perspective of human operators, the requirement to move the periscope also introduces the need to store information in memory, solve motion artefacts, and integrate visual information over time (Burr, & Cicchini, 2014). These perceptual and cognitive requirements could exceed available operator resources, thereby impairing task performance.

As navies shift to digital technologies, however, the options for presenting visual information become flexible. A digital periscope apparatus is capable of representing a wider field of view (FOV - up to the full 360°), thus potentially accelerating the detection of contacts, and reducing perceptual and cognitive requirements arising from the narrow FOV of a traditional physical periscope. Moreover, a long history of psychology research assessing

visual search behaviour demonstrates that humans are highly efficient at detecting stimuli that possess unique attributes that differ from their environment, such as colour or shape (Treisman & Gelade, 1980). Such *salient* stimuli tend to guide attention (Egeth, 1977) efficiently to their location, independent of observer attentional set (Posner & Cohen, 1984; Wolfe et al., 2003), thus capitalizing on what has been argued to be the default strategy for directing attention to the environment (Eckstein, 2011; Kohler et al. 2014).

On the other hand, a larger FOV display could create cognitive and perceptual demands whose impacts on performance are currently not understood. The requirement for operators to scan more visual information simultaneously raises several possible problems. First, practical limits on display size imposed by physical space constraints in the submarine mean that a contact presented on the 360° digital periscope display will be physically smaller than the same contact presented on a display showing a narrow rotatable view. Thus, contact image details may be harder to discern, potentially impairing contact detection speed and/or accuracy. Second, a larger FOV will contain more information, potentially introducing visual clutter (Moacdieh & Sarter, 2015), interference (Ramachandran & Cobb, 1995; Whitney & Levi, 2011) and response competition (Eriksen & Eriksen, 1974; Reddy & VanRullen, 2007) that could slow contact detection, and/or induce search fatigue.

In summary, while a digital periscope brings with it the opportunity for larger FOVs, the visual properties of these display could negate their potential benefits of increased information availability and reduced operator load. This uncertainty highlights the need for systematic research to evaluate the impact of digital periscope HMI concepts on operator performance. To this end, we compared contact detection performance using two alternate HMI concepts in a representative safety-search style simulated submarine mission.

Experiment 1

Two simulated digital periscope HMI concepts were developed by the Australian Defence Science and Technology Group as prototypes. The HMI designs emerged from work with the Royal Australian Navy, using cognitive work analyses (CWA) and human-centred design workshops.

The first concept, called OPTIX, displayed a single-bearing view of the environment consisting of an arc of the horizon. The visual view was rotated horizontally across the horizon by a fixed seated participant via keyboard (arrow keys) or mouse (click and drag horizontally), simulating a conventional periscope. The second concept, called OPTIMUS (OPTronics IMplementation & Usability System), displayed a full 360° environment across five panels, presented on a single monitor identical in size to the one used in OPTIX.

Each concept display operated at low or high magnification. For OPTIX, the display showed an arc of 36° of the horizon on low magnification and an arc of 3° of the horizon at high magnification (Figure 1). For OPTIMUS, each panel displayed an arc of 75° of the horizon (there was slight overlap over each panel; Figure 2a) on low magnification. At high magnification, the mouse could be used to move a magnified window to inspect a bearing section on any panel on OPTIMUS. This magnified window showed an arc of 3° of the horizon at its location (see Figure 2b). Thus, at low magnification, contacts in OPTIMUS were approximately half the size of contacts in OPTIX, while at high magnification contacts in both HMIs were of equivalent size.

(a)

(b)

Figure 1. OPTIX display. A Warship at a range of 9 kyd at (a) low magnification (36° horizon arc), and at (b) 12x high magnification (3° horizon arc).

(a)

(b)

Figure 2. OPTIMUS panorama display showing the full 360° horizon in five panels. A Warship at a range of 9 kyd (a) without the use of the magnified window, and (b) with the use of the 12x magnified window.

Experiment 1 compared contact detection, perceived system useability and perceived workload across OPTIX and OPTIMUS in scenarios with continuously-appearing contacts that required active monitoring of the periscope displays. We manipulated contact distance by presenting half the contacts “Near” (average onset range 3.5 kyd) and the other half “Far” (average onset range 13.8 kyd) from Ownship. The distance manipulation affected the size of each contact in terms of its visual angle and the amount of the contact visible above horizon (e.g. mast only). However, other factors that impact on the visibility of far-range contacts (e.g. haze, light scatter) were not modelled, and for this reason the imagery of both Near and Far contacts was more similar to real-world short-range contacts. Participants completed scenarios using low magnification displays, high magnification displays (or the magnification window for OPTIMUS), or “hybrid” magnification in which participants could choose when to alternate between high and low magnification.

Method

Participants

Twenty-four participants (10 males and 14 females) were recruited from The University of Western Australia (UWA). Participants took part for compensation of \$20 AUD. The mean age of participants was 27.2 ($SD = 8.7$). This research complied with the American Psychological Association Code of Ethics and was approved by the Human Research Ethics Office at UWA (RA/4/20/5883). Informed consent was obtained from each participant in person prior to the experiment.

HMI Concepts and Stimuli

Each HMI was displayed on a 1920 x 1080, 27-inch display (Figures 1 and 2). Imagery was created in the Unity graphics engine and the HMI was coded in Python. The same Unity models (contacts) were displayed in each experimental HMI. Both HMI concepts expressed range in yards and kiloyards (kyd) to ensure consistency with submarine systems

(1000 yards = 914 meters). For OPTIX, in the hybrid magnification condition the FOV could be changed between these two extremes, using the plus and minus keys or the slider on the display. For OPTIMUS, in the hybrid magnification condition, the magnified window could be toggled on and off by holding down or releasing the mouse button. Participants marked detected contacts by double-clicking on them. In both HMIs, a marker was added to the bearing strip when a contact was detected. The time at which each contact was detected was recorded by the system. A contact was recorded as detected if the participant double-clicked within 10 degrees either side of the contact. Any contact that was not “detected” by the end of the scenario was considered “missed”.

Participants were instructed how to conduct the detection task under each magnification condition. For OPTIX, participants scanned only in low magnification when in the low magnification condition, and used only high magnification when in the high magnification condition. In the hybrid magnification OPTIX condition, participants were encouraged to start with a quick low-magnification all-round scan before using high magnification. For OPTIMUS, participants in the low magnification condition were instructed to search for contacts as preferred. Participants in the high magnification condition were encouraged to systematically scan the entire panorama with the high magnification window and to respond to a potential contact only when they reached it in their scan pattern. This was designed to make this condition comparable to the high magnification condition in OPTIX where only contacts in the viewing arc could be marked. In the hybrid magnification condition, participants were encouraged to do a quick sweep of the panorama without magnification at the beginning of each scenario to detect immediately visible contacts.

Contact Detection Scenarios

We developed six scenarios, each with 20 contacts, that lasted six minutes each. Each scenario had the same contact types (e.g. warships, fishing vessels) that differed in visual

characteristics (e.g., colour, height length). Although these contacts features were consistent across each scenario they appeared at different bearings to counter learning effects. Contacts appeared at 18° spans around Ownship, with the bearing of the contact chosen randomly except that it was at least 6° from the edge of its span. Contact bearings remained constant for the duration of the scenario.

Each participant completed OPTIX and OPTIMUS scenarios at all three magnification conditions. OPTIMUS and OPTIX conditions were completed in blocks, with the order of magnification determined using a Latin square. Each scenario was assigned to each HMI and each magnification at least twice across participants. Participants were randomly assigned to condition order. Contacts were evenly divided between “Near” and “Far”. Far contacts gradually came over the horizon towards the Ownship. Eight Far contacts could be detected from the start of the scenario (four were visible, as assessed by experimenters, under low magnification and four required high magnification), while two became visible later in the scenario. Three Near contacts were visible from the start of the scenario, while the remaining seven abruptly appeared at intervals throughout the scenario (as might be expected to occur if a contact had emerged from behind another contact). Each contact remained on the screen for the duration of the scenario.

Measures

We recorded detection accuracy and time-to-detect contacts from the time they were visible. This primarily reflects the safety aspect of a periscope operators’ role – identifying contacts quickly and accurately is essential for stealthy submarine operation. To broaden our assessment of the impact of the systems, we also measured perceived workload and perceived system usability. We used the Modified Cooper Harper (MCH: Wierwille & Casali, 1983) scale to assess mental workload. The MCH uses a decision tree to elicit operator perceived workload, on a scale of 1 to 10. Usability was measured using the System Usability Scale

(SUS; Brooke, 1996). Example questions include “I found the various functions in this console were well integrated” and “I found this console very cumbersome/awkward to use.”

Procedure

Experimental sessions began with verbal instructions outlining task requirements. Participants were then introduced to the MCH and trained to use OPTIMUS or OPTIX, depending on to which they were first assigned. All three magnifications were demonstrated, and practice was conducted with hybrid magnification. As the hybrid condition allowed the use of both low and high magnification, participants were able to practice each magnification.

Following training, each participant completed three scenarios with their first HMI, using a different search strategy (low, high, and hybrid) for each scenario. Prior to beginning each scenario, participants were reminded of the magnification requirements.

After finishing the first three scenarios, participants completed the workload and usability questionnaires, and then were trained on the other HMI and repeated the experimental procedure.

Results

Analysis Approach

We report Bayesian statistics. Quantifying uncertainty amongst competing hypotheses (including the null) using Bayesian statistics is a principled and well supported practice (Vandekerckhove et al. 2018) and avoids some of the critiqued limitations of frequentist statistics (Wasserstein & Lazar, 2016, Wagenmakers et al, 2018a). We report the *posterior probability* that a model (e.g., ANOVA with a specific combination of terms) generated the observed data, given the true model is in the set tested (i.e., the likelihood of a specific combination of parameters being the “true model”). In place of *p*-values, we report Bayes Factors (BF), which can be interpreted as the strength of evidence for one hypothesis over another (e.g. a Bayes Factor of 10 means one hypothesis is 10 times more likely than the

other; Wagenmakers et al, 2018b). We follow the commonly reported standard of $BF > 3$ as “weak evidence”, $BF > 5$ as “moderate evidence”, $BF > 10$ as “strong evidence”, and $BF > 100$ as “very strong evidence” (e.g. Etz & Vandekerckhove, 2016), avoiding the use of a single arbitrary cut-off for our hypothesis tests.

Another advantage of Bayesian statistics is the ability to quantify evidence *for the null* (as opposed to simply failing to reject it). To distinguish evidence for the Alternate vs Null hypotheses we report BF_{10} when evidence favours the alternate hypothesis (i.e., that there *is* a difference or effect), and BF_{01} when evidence favours the null hypothesis of no difference. When reporting descriptive statistics we report the 95% Bayesian Credible Interval (BCI: Wagenmakers et al., 2018b) in line with the Bayesian statistical framework used throughout. All analyses were conducted using the free statistical software JASP (Love et al, 2019).

Data Exclusion

With 24 participants detecting 20 contacts in each of six scenarios (2 HMI x 3 Magnification), there were a total of 2,880 detection events. However, the final contact, presented approximately 50 seconds prior to the end of the scenario was detected only 54.2% of the time compared with an average hit-rate of 96.5% for other contacts. This may have been due to the limited time participants had to detect the contact. On this basis, we excluded this contact, leaving 2,736 detection events for analysis.

Contact Detection Accuracy

Accurate detection was defined as clicking within a 10° bearing of a contact and was calculated as the proportion of the total number of contacts detected in each scenario.

Accuracy as a function of HMI, Magnification and Distance is presented in Figure 3. A Bayesian ANOVA on these data, controlling for subject-level variance with a random “Subject” effect, favoured a model including Magnification, Distance, and their interaction (posterior probability = 0.95). There was strong evidence *against* an effect of HMI

($BF_{01}=54.01$) on contact detection accuracy. Conversely, there was very strong evidence *for* an effect of Magnification on accuracy ($BF_{10}>1000$). High and Hybrid magnification did not show any accuracy differences ($BF_{01}=29.10$), but both had greater accuracy than Low magnification ($BF_{10}=224.18$, and $BF_{10}>1000$ for High and Hybrid respectively). The very strong evidence for an interaction between Magnification and Distance ($BF_{10}>1000$) reflects the fact that accuracy approached ceiling levels for Near contacts, regardless of Magnification, while accuracy was poorer for Far contacts with Low magnification compared to High and Hybrid magnification.

Figure 3. Detection Accuracy for contacts in Experiment 1 by Magnification and Distance. Collapsed across OPTIX and OPTIMUS. Error bars reflect the 95% Credible Interval around the mean.

Contact Detection Time

We computed detection time as the difference between the time a contact was first visible (onset time) and the time it was detected (detect time). Onset time was defined as the moment the contact was discernible on the horizon to an operator looking directly at its

location (determined by the experimenters looking for known single contacts), and was calculated separately for each HMI and magnification.

Detection time as a function of HMI, Magnification and Distance is presented in Figure 4. A three-way Bayesian ANOVA on these data favoured a model including all three main effects, plus the HMI x Magnification and Magnification x Distance interactions (posterior probability=0.883). There was very strong evidence for the inclusion of HMI, Distance and Magnification main effects, and the interactions of Magnification x Distance, and HMI x Magnification (all $BF_{10} > 1000$). There was strong evidence against the interaction of HMI x Distance ($BF_{01} = 14.88$) but inconclusive evidence for or against the three-way interaction ($BF_{01} = 1.04$).

The OPTIMUS panorama view resulted in faster contact detection than OPTIX. Low Magnification was faster than Hybrid, which was faster than High. Near contacts were detected faster than Far contacts. The HMI x Magnification interaction reflects the smaller difference in contact detection time across HMIs at High magnification than at either Low or Hybrid (Figure 4, left panel). While JASP cannot compute Bayes Factors for post-hoc interactions, Tukey-corrected p-values showed significantly faster detection for OPTIMUS compared with OPTIX at Low ($p_{Tukey} = 0.005$) and Hybrid ($p_{Tukey} = 0.002$), but not at High ($p_{Tukey} = 1.00$) magnification. The Magnification x Distance interaction was such that the Magnification effect was flipped for Near contacts versus Far contacts. With Near contacts, Low and Hybrid magnification provided a detection time advantage for both HMIs (in OPTIMUS the Low magnification was statistically faster than Hybrid). For Far contacts, Low and Hybrid magnifications slowed detection times for OPTIX, but for OPTIMUS the Hybrid magnification facilitated the fastest responses.

Figure 4. Detection Time to detection for Near (left panel) and Far (right panel) contacts, by HMI and Magnification. Error Bars represent the 95% Credible Interval around the mean.

Perceived Workload and Usability

On the MCH, a higher score (out of 10) indicates higher reported workload. A Bayesian paired-samples *t*-test showed inconclusive evidence for the Null hypothesis ($BF_{01}=2.42$) indicating that MCH scores for OPTIMUS ($M=3.26$, 95% CI [2.05, 4.78]) and OPTIX ($M=3.78$, 95% CI [2.23, 4.77]) were indistinguishable. In practice, participant rating varied wildly, with both systems receiving scores from 1 to 9. On the SUS, a higher score (out of 10) indicates a more usable system. Three participants failed to complete one question. We thus scored their results out of 9, and then re-scaled them to be in the same range as the other participants. A Bayesian paired-samples *t*-test showed no evidence for or against a difference in the perceived usability between OPTIMUS ($M=77.41$, 95% CI [71.60, 83.22]) and OPTIX ($M=71.42$, 95% CI [64.35, 78.50], $BF_{01}=2.14$).

Discussion

Experiment 1 compared contact detection between OPTIX and OPTIMUS across Near and Far contacts using three magnifications. Overall, there was no difference in detection accuracy between the HMI concepts. However, OPTIMUS did yield faster detection of contacts compared with OPTIX. The detection time advantage for OPTIMUS compared to OPTIX was exhibited under Low and Hybrid magnification strategies, but was absent with High magnification where both systems had restricted 3° FOVs. This interaction points toward the effectiveness of OPTIMUS using low and hybrid strategies in attracting a more efficient search process as the panorama allows attention to be guided to salient stimuli within the operator's peripheral vision, and the High magnification condition specifically prevents this from occurring in both HMIs, resulting in the exhibited OPTIMUS induced performance equivalent to the serial OPTIX HMI. The OPTIMUS detection speed advantage held for both Near and Far contacts. The pattern of results for Far contacts specifically highlighted the combined benefit of the panoramic display with the hybrid strategy, where it

is likely that some contacts immediately ‘popped out’, while others were small enough that they required magnification to be identified properly. In this way, the hybrid panoramic (OPTIMUS) system seems best tailored to the human visual search system, which combines fast, broad coverage searches (Wolfe, 2021) with more guided, spotlight style searches (Cave & Bichot, 1999). That is, the OPTIMUS panoramic system provides access to all visual information at once, allowing the efficient automatic search for maximal speed advantage, and the hybrid magnification tool facilitates the guided spotlight search component.

Overall, our data suggest that the disadvantage of having smaller sized contacts displayed in the OPTIMUS panorama view compared to OPTIX at low magnification was outweighed by the advantage of having continuous access to the 360° visual horizon in OPTIMUS, and could be further overcome by access to operator-guided magnification tools. There was no difference between HMIs in terms of perceived workload or system usability. We conclude that for the simulated submarine periscope task of contact detection under high visibility/contrast conditions, the OPTIMUS panorama view offers advantage without any accompanying costs.

Experiment 2

In addition to the differences in its representation of the external world, a digital periscope provides opportunities to add digital tools to aid operators in tasks such as range estimation (the distance between Ownship and the contact) and course estimation (inferred from contact visual aspect or ‘angle to the bow’ - ATB). While, on the face of it, digital aids should be superior to conventional manual estimation techniques, the increased information presentation associated using digital tools and potential associated display clutter/interference/response competition (Eriksen & Eriksen, 1974; Moacdieh & Sarter, 2015; Whitney & Levi, 2011), could increase cognitive load, and this could outweigh potential advantages. With this in mind, we examined whether complex tasks are better

supported by the digital tools available in OPTIMUS compared to conventional manual aids used in with OPTIX, and whether digital tools influenced perceived workload and useability.

To avoid confounding the effects of digital tools and differences in visual displays (as shown in Experiment 1) between OPTIMUS and OPTIX, all contacts were automatically detected and highlighted for participants in Experiment 2. Contact classification was then undertaken for each highlighted contact using tools that were common to both OPTIMUS and OPTIX (with minor differences in the physical location of the tools on the display). Contact range and course estimates were differentially aided by digital tools across the HMIs. Digital 3D models of potential contacts in the analysis pane were used in OPTIMUS (Figure 6C), which allowed participants to match digital object model features directly to contact images to allow for more accurate contact range and course estimates. By comparison, in OPTIX, participants estimated contact range using a combination of heuristics, system calculations, and a physical model akin to procedures used with conventional periscopes (Figure 7).

Finally, we also tested another potential feature of future digital periscopes – automated contact tracking. For both OPTIMUS and OPTIX, in half of scenarios, contacts were “auto-tracked” (the periscope view automatically followed a contact so it did not drift out of the analysis pane as it moved relative to Ownship), while in the other half, participants had to manually track contacts.

Method

Participants

Sixteen (10 female, 6 male) participants were recruited from the UWA. Participants took part for compensation of \$75 AUD. Mean age of participants was 24.8 years ($SD = 5.0$).

HMI Concepts and Stimuli

The bottom screen of HMIs provided a range of information panels and tools, the layout and sizes of which were driven by the outcomes of a CWA process with operational

submariners (Schmitt et al., 2013; Irons et al., 2019). As a result, the information provided by both OPTIX and OPTIMUS for the bottom panel was identical, although their layout and sizes had some differences (see Figures 5 and 6 for the features of each HMI design).

Participants completed two scenarios using each HMI (with and without auto-tracking) in counterbalanced order. In OPTIX, range was estimated by aligning the bars of a range estimate tool (Figure 5B) with reference to lines on a standard 2D broadside silhouette of that class of contact in the classification database (Figure 5F) and selected a corresponding height value from the 2D silhouette to generate a range estimate. If the contact was partly obscured by the horizon and/or a height range was chosen on the contact that was not available on the silhouette, the participant needed to add or subtract height values to match the chosen height range. To estimate ATB, participants rotated the manual ATB model (Figure 7) until it matched the direction of the contact on their display.

In OPTIMUS, range was estimated by positioning the top and bottom lines of a ranging box (Figure 6B) over an image of the contact in the analysis panel, and then positioning a second set of lines to corresponding locations on the 3D contact model shown in the range/ATB tool (Figure 6C). Thus, for OPTIMUS, there was no need to fit a specific height measurement as this was done by the system, provided the contact had been classified. ATB was estimated by adjusting the angle of a 3D model using a slider until it matched the visible appearance of the contact. Both range and ATB fields were populated in the range and ATB window.

For both HMIs, participants manually entered the range and course estimates as a single entry per contact. This entry was referred to as a ‘visual cut’ and simulated a periscope operator feeding visual estimates to team members. To enter information, participants pressed the “cut” button on the range and ATB window. Participants were asked to submit one cut per contact, and repeat cuts were excluded from analysis.

Figure 5. The OPTIX display as presented on two vertically aligned screens, with key features marked. A: Visual cut window. Range information could be sent to this window once operators had manipulated the range bars (B), and ATB (course) estimates were entered manually. B: Range bars, which operators adjusted to match the lines on the image as displayed in the classification database (F). C: Contact height bins specific to the contact classification. A height bin was selected with reference to the contact's classification image in the classification database (F) and position of range bars (B). D: Magnification slider for controlling magnification. E: Contact tote listing the contacts currently in the system and their associated data. F: Classification database with height intelligence. Images of different classification types could be viewed in this database, allowing participants to match a visible contact to a classification type. Once a classified contact was selected for analysis, the classification database provided height reference lines for ranging. G: Geoplot showing a bird's eye view of the location and direction of travel of Ownship submarine (centre) and surrounding contacts. Unlabelled HMI elements were not referred to in this experiment.

Figure 6. The OPTIMUS display as presented on two vertically aligned screens, with key features marked. A: Visual cut window. The range and ATB (course) information in this window automatically updated as operators manipulated the yellow ranging box (B) and range/ATB tool (C). B: Ranging box in the analysis panel. Operators adjusted the position of the top and bottom lines to self-selected positions on the contact, and then matched these positions by manipulating the ranging box in the range/ATB tool (C) to generate a range estimate. C: Range/ATB tool showing a 3D model of the currently selected contact's classification. In addition to the ranging box described above, the slider at the bottom of this tool could be dragged to rotate the 3D model to match the visible contact. This generated an ATB estimate which automatically populated the visual cut window. The range/ATB tool could be toggled to the classification database in order to classify contacts. D: Magnification slider for controlling magnification in the analysis window. E: Contact tote listing the contacts currently in the system and their associated data. F: Classification tab, which switched the Range/ATB tool to the classification database (identical to F in Figure 5). The panel could be switched back to the Range/ATB tool using the "Ranging" tab. G: Geplot showing the location and direction of travel of Ownship submarine (centre) and surrounding contacts. H: Magnified window, which could be controlled by mouse to inspect sections of the panorama. Unlabelled HMI elements were not referred to in this experiment.

Figure 7. The manual ATB tool used by OPTIX operators. Operators rotated the tool to match the orientation of a contact on their display and noted the corresponding ATB estimate.

Scenarios

Two scenarios were generated, each with ten contacts that were automatically marked on the bearing strip and geoplot. At the start of the scenario, contacts were positioned at evenly-spaced bearings, with the classification at each bearing randomised. The speed of contacts was held constant at 5 knots (about 9 km/hr). Contact courses were set so that contacts of the same type had one port ATB and one starboard ATB, one was closing (heading towards Ownship) and the other opening (heading away), and the average ATB angle was the same across scenarios.

A further two scenarios were generated by flipping each of the two original scenarios along a North-South line (passing through Ownship) to create mirror image versions. Scenarios were assigned such that participants did not complete a pair of mirrored scenarios on the same HMI design.

Measures

We recorded classification accuracy and estimated classification speed as the mean time interval between each classification entry (“inter-classification time”). Range estimation error was quantified as the absolute deviation of the visual cut range from the simulation truth range in yards. Because range estimation was correlated with true range (Pearson’s $r=0.313$, 95% CI [0.239, 0.383], $BF_{10}>1000$), we divided it by true range to create a “percentage range error” that was uncorrelated with true range. Course estimation error was quantified as the absolute deviation of the visual cut course from the simulation truth course, expressed in degrees. The speed of range and course estimation was quantified as the mean time interval between each visual cut (“inter-cut time”). Perceived workload and usability were measured as in Experiment 1.

Procedure

Experimental sessions began with a twelve-minute video introducing participants to the periscope role, relevant nautical concepts such as port and starboard, and the tasks of classification, contact prioritisation, and contact range and ATB estimation. Participants were then trained to use the MCH and introduced to either OPTIMUS or OPTIX (order counterbalanced) and were trained to perform their tasks (with auto-tracking disabled).

Contacts were numbered sequentially and participants were instructed to classify each contact in numerical order. This was achieved in both HMIs by matching each contact with a contact type found in the classification database and updating the contact information in the contact tote (Figures 5E and 6E). Classification automatically changed the contact icon in the contact tote to a fishing, merchant or warship icon, allowing a contact’s classification to be found at a glance on the geoplot (Figures 5G and 6G). Once all contacts had been classified, participants began estimating their range and ATB. Participants were instructed to complete estimates in order of priority (warships followed by fishing boats followed by merchants). Within each contact type, participants were instructed to prioritise closer contacts.

Following training with one HMI, participants completed two scenarios using that HMI with auto-tracking on or off (counterbalanced). Before each scenario, participants were informed that they would either have to manually adjust the viewing window to keep contacts in view (no auto-tracking), or that the system would automatically keep the contact in view (auto-tracking). Scenarios continued until participants had submitted visual cuts for all 10 contacts, which took an average of approximately 15 minutes. Following completion of the first two scenarios, participants completed the battery of questionnaires.

After a short break, participants were trained to use the other HMI and repeated the experimental procedure with that design.

Results

Contact Classification

Accuracy. Over four scenarios, participants provided 640 classification data points in total (16 participants x 4 scenarios x 10 contacts). Overall accuracy was high (93.1% in OPTIMUS; 96.5% in OPTIX), and a two-way Bayesian ANOVA showed no evidence that either Tracking ($BF_{01}=16.07$) or HMI ($BF_{01}=2.67$) influenced classification accuracy.

Speed. A two-way Bayesian ANOVA on inter-classification time preferred the model including HMI only (posterior probability = 0.86). There was strong evidence for a main effect of HMI ($BF_{10}=23.29$), and against the inclusion of Tracking ($BF_{01}=12.08$) and the interaction ($BF_{01}=20.09$). Classification was reliably faster in OPTIX than OPTIMUS (Figure 8).

Figure 8. Classification response time for each HMI for autotracking present and not present. Error Bars represent the 95% Credible Interval around the mean.

Range Error

Contacts that were mis-classified were excluded from analysis of range and course data (leaving 603 valid data points for each). A Bayesian ANOVA on the range percentage error by HMI and Tracking suggested that the model only including HMI was preferred (posterior probability = 0.84). There was very strong evidence for the main effect of HMI ($BF_{10} > 1000$). Range error was approximately half as large in OPTIMUS as it was in OPTIX (Figure 9). There was moderate evidence against both Tracking ($BF_{01} = 7.69$) and the interaction term ($BF_{01} = 9.80$).

Figure 9. Range error as percentage of range for each HMI for autotracking present and not present. Error Bars represent the 95% Credible Interval around the mean.

Course (ATB) Error

A Bayesian ANOVA on course error data by HMI and Tracking indicated that the model including only HMI was preferred (posterior probability >0.99). There was very strong evidence for a main effect of HMI ($BF_{10} > 1000$). Course error was approximately half as large in OPTIMUS as it was in OPTIX (Figure 10). There was strong evidence against the main effect of Tracking ($BF_{01} = 10.332$) and the interaction term ($BF_{01} = 17.11$).

Figure 10. Course error in degrees for each HMI for autotracking present and not present. Error Bars represent the 95% Credible Interval around the mean.

Course and Range Estimate Speed

A Bayesian ANOVA on inter-cut time by HMI and Tracking preferred the null model (posterior probability = 0.544). Consistent with Figure 11, there was uncertain evidence for or against the main effect of HMI ($BF_{01}=2.3$), and strong evidence against the inclusion of Tracking ($BF_{01}=12.69$) and the interaction ($BF_{01}=42.79$).

Figure 11. Time to generate solutions (in seconds) for each HMI for auto-tracking present and not present. Error Bars represent the 95% Credible Interval around the mean. Note, there may be noise in the measurement if participants operate contrary to the order instructed (e.g., computing a solution before classifying all contacts). Increased error bars may reflect violations of this behaviour by some participants.

Perceived Workload and Usability

We collected MCH ratings for each HMI separately as a function of the presence/absence of auto-tracking. A Bayesian ANOVA on MCH rating by HMI and Tracking preferred the model including both main effects (posterior probability = 0.538). There was very strong evidence for the inclusion of HMI ($BF_{10} > 1000$), but uncertain evidence for the inclusion of Tracking ($BF_{10} = 1.72$). MCH ratings were lower for OPTIMUS than OPTIX, and lower when auto-tracking was present compared to absent (Figure 12). We collected SUS ratings once per HMI (collapsed across presence/absence of auto-tracking). Three participants failed to complete a question as outlined in Experiment 1, and thus SUS scores for these participants were rescaled as per Experiment 1. A Bayesian paired samples t-test showed weak evidence that OPTIMUS ($M = 78.91$, 95% CI [70.93, 86.87]) was rated as more usable than OPTIX ($M = 61.64$, 95% CI [51.33, 71.96], $BF_{10} = 4.18$).

Figure 12. Mean MCH ratings by HMI and Auto-Tracking. Error Bars represent the 95% Credible Interval around the mean.

Discussion

In addition to the differences in its representation of the external world, a digital periscope provides opportunities to add digital tools to aid operators in the tasks such as range estimation and course estimation. We examined whether these tasks are better supported by the digital tools available in OPTIMUS compared to conventional manual aids used in with OPTIX, and whether digital tools influenced perceived workload and useability.

Experiment 2 compared contact classification, range, and course estimation between an OPTIMUS concept that provided digital tools to assist with range and course estimation, and an OPTIX concept that provided limited decision aids in keeping with conventional non-digital periscope systems. Course and range estimates were approximately 2x better when using OPTIMUS compared to OPTIX. However, we found no reliable evidence that HMI impacted the time taken to make these estimates. This suggests that the digital tool improved the efficiency of evaluating contact behaviour (as participants generated more accurate solutions in comparable time-frames). In addition, participants perceived workload to be lower when using OPTIMUS compared with OPTIX. There were no differences in

classification accuracy across the HMIs, suggesting that advantages in range and course estimation in OPTIMUS cannot be attributed to any differences in classification accuracy. In fact, participants took longer to classify contacts when using OPTIMUS than OPTIX. Although the classification tools themselves were identical apart from minor location changes, this may reflect a potentially greater number of clicks to navigate to the OPTIMUS classification pane

We also found that auto-tracking did not provide reliable advantage to either HMI. We caution that this interpretation may be dependent on the fact that our contacts moved relatively slowly and were highly visible. Tracking faster-moving contacts (particularly when closer to Ownship) and less-visible contacts would potentially benefit more from reliable auto-tracking. These possibilities warrant further investigation.

General Discussion

In the near future, digital periscope concepts may offer 360° FOV and enable digital tools to aid submarine periscope operators in detecting and tracking contacts. However, questions remain about whether the potential benefits might be outweighed by the burden they place on operator perceptual and cognitive resources. To address these questions, we separately examined the impact of increasing the periscope FOV on contact detection (Experiment 1) and the impact of providing digital tools on estimating contact course and range (Experiment 2) by comparing performance on a digital HMI concept (OPTIMUS) and a concept that simulated a conventional manual periscope (OPTIX).

Contact Detection

In Experiment 1 contact detection was faster in OPTIMUS than OPTIX. This indicates that although the physical size of contacts on the display was equal (at high magnification) or smaller (at low magnification) in OPTIMUS, providing a 360° FOV speeded contact detection compared to a narrow rotatable FOV. Moreover, there was no

evidence that OPTIMUS negatively impacted perceived workload or system usability compared to OPTIX.

Overall, the results suggest that a wider FOV may allow operators to more efficiently allocate attention to contacts, and that the additional information in these displays does not negatively impact operator perceptual and cognitive resources. That said, factors such as near sea-surface atmospheric conditions, weather, and lighting (e.g. time of day) impact the visibility and salience of contacts and thus visual search performance (Hogervorst et al., 2013). For this reason, future research should examine whether wider FOV displays, like those in OPTIMUS, are still superior under reduced visibility conditions.

One possibility, suggested by our results, is to train operators on hybrid magnification, thus ensuring the continued search advantage of OPTIMUS panorama view (e.g., the most salient contacts are acknowledged first, then a serial search pattern is performed to detect less immediately salient contacts). This strategy maps well to theoretical accounts of guided visual attention (Wolfe, 2021) and can help overcome performance decrements when magnification was not available for the higher FOV system. With the hybrid strategy OPTIMUS provides all the benefits of immediate, automatic attention strategies and facilitates the same serial, guided search as conventional periscopes while maintaining the panoramic view in peripheral vision.

Digital Tool Support

In Experiment 2, we found high classification accuracy, but that participants took longer to classify contacts when using OPTIMUS, potentially caused from needing at least an extra mouse click to access the classification panel if it were behind the ATB/ranging tool (a practical requirement borne of needing extra screen space to display a magnified image of the selected contact in the panorama).

We also found that OPTIMUS, which provided participants a suite of digital tools that might be plausible in future periscope systems, resulted in benefits to course and range estimate accuracy, with smaller error (approximately halved) compared to OPTIX. We tentatively suggest this stems from the fact that our tools provided 3D digital models of contact types, enabling operators to use a feature-matching comparison to estimate contact course and range. For ranging, this meant participants could match the features on the model they chose to anchor digital ranging lines with the identical features on the detected contact. Previous research has shown that matching to a known sample produces finer/more accurate estimates than making absolute judgments of a single visual object (Bell et al., 2013). With respect to course estimates, previous studies suggest that 3D representations can improve perceptual accuracy relative to 2D representations like those that were used in OPTIX (Bell et al., 2015). Further, there was no need to consider actual heights or mentally subtract or add these heights under some viewing conditions to calculate the range as was the case in OPTIX. Indeed, participants reported lower workload when using OPTIMUS, suggesting the 3D models made it easier to estimate contact range and course.

While Experiment 2 suggests clear performance advantages arise from digital tools in the short term, this support could come at a cost, including complacency and associated loss of situation awareness (Parasuraman, & Manzey, 2010) or skill degradation (Casner et al., 2014). For example, in the longer term, matching moveable lines on the display with the 3D tool, with little to no reference to the physical height of the contacts themselves may interfere with the development of operator expertise. With current systems, operators become increasingly familiar with the heights of certain contacts through experience which can facilitate range estimates for familiar contacts and serve as an anchor to estimate the range of less familiar contacts. In the absence of this experience, operators may become reliant on the HMI tools, impairing their ability to detect potential errors or failures in digital tool

operation, and unable to efficiently estimate contact course and range without the tools operating reliably. Thus, while digital tools, like those in OPTIMUS, could provide performance benefits, future research should examine potential costs that could arise.

Study Limitations

This study used a medium-fidelity simulation with imagery quality and contact models matched across the two HMIs. As noted earlier, one limitation is the near-ideal visibility and sea state conditions in our HMIs. The contrast between contacts and their surroundings is often much lower (e.g. when sea state, weather or lighting conditions are poorer). Such conditions would likely increase visual clutter and reduce target salience, thus impairing search efficiency (Duncan & Humphreys, 1989; Liesefeld et al., 2016). In addition, the visual scenes in the HMI were relatively sparse with low contact similarity and density. If contact similarity were higher (i.e., more similar contacts) and density were increased (i.e., more adjacent contacts with similar features), the search advantages afforded by wider FOV might decrease. On the other hand, to the extent that operators could continue to function efficiently under these kinds of sub-optimal viewing conditions, digital tools like those in OPTIMUS might be even more useful than demonstrated in our research.

Another limitation of our work is that the impact of increased FOV and digital tools were examined separately in order to avoid confounding their effects on performance. In reality, operators continue to monitor the periscope for contacts even whilst analysis is undertaken on a priority basis, and the various types of periscope operator tasking overlap and are interdependent. Thus, there is a need to replicate this work under simulated conditions more akin to operational conditions and with participants undertaking tasking that more realistically mimics that of an actual periscope watchkeeper.

Finally, in addition to the need to conduct a more complete evaluation of the potential advantages of auto-tracking capability noted earlier, additional work is needed to examine the

impact of digital periscope concepts on situation awareness. Although contact detection and analysis performance may benefit from the shift to digital periscopes, there may be a cost to situation awareness arising from dependence on digital tools akin to those arising from dependence on automation (Onnasch et al., 2014).

Conclusion

We demonstrated the potential performance advantage afforded by a digitally supported panoramic field of view, compared to maintaining a single, narrow view display used in conventional periscope technologies. Our OPTIMUS HMI concept showed robust advantages for contact detection, and range and course estimates at least with the tasks and viewing conditions simulated. Our work demonstrates the potential benefits of considering psychological insights into human perception and cognition in designing future HMIs. However, future research will need to consider the complex operational conditions under in which HMIs must function effectively as well as the potential costs of digitally supported panoramic FOVs such as operator complacency and skill degradation.

References

- Bell, J., Kanji, J., & Kingdom, F. A. (2013). Discrimination of rotated-in-depth curves is facilitated by stereoscopic cues, but curvature is not tuned for stereoscopic rotation-in-depth. *Vision Research*, 77, 14-20.
- Bell, J., Manson, A., Edwards, M., & Meso, A. I. (2015). Numerosity and density judgments: Biases for area but not for volume. *Journal of Vision*, 15(2). doi:10.1167/15.2.18
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London: Taylor & Francis.
- Burr, D., & Cicchini, G. M. (2014). Vision: efficient adaptive coding. *Current Biology*, 24(22), R1096-1098.
- Casner, S. M., Geven, R. W., Recker, M. P., & Schoolar, J. W (2014). The retention of manual flying skills in the automated cockpit. *Human Factors*, 56, 1506-1516.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, 6, 204-223.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of vision*, 11(5), 14-14.
- Egeth, H. (1977). Attention and preattention. *Psychology of Learning and Motivation*, 11, 277-320, Academic Press.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143-149.
- Etz, A., & Vandekerckhove, J. (2016). A Bayesian perspective on the reproducibility project: Psychology. *PloS one*, 11, e0149794.

- Hogervorst, M., A., Toet, A., & Bijl, P. (2013). Human search with a limited field of view: effects of scan speed, aperture size, and target conspicuity. *Optical Engineering*, 52, 041106.
- Irons, J., Schmitt, M., Michailovs, S., & Huf, S. (2019, September 2-5). *Human performance in the maritime control room: Using simulation for exploring team cognition and information layer design* [Paper presentation]. Simulation Congress 2019, Gold Coast, Australia. <http://www.simulationcongress.com/about-2/history/asc2019/>
- Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 821-836.
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., ... & Wagenmakers, E. J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88, 1-17.
- Moacdieh N, Sarter N. (2015). Display clutter: A review of definitions and measurement techniques. *Human Factors*, 57, 61-100.
- Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, 56, 476–488.
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52, 381–410.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes*, 32, 531-556.

- Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature*, 373, 66-68.
- Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision*, 7, 3-3.
- Schmitt, M. L., Parera, D. L., Henley, P. J., & Myer, A. (2013). *A Cognitive Work Analysis-based user interface design process: A conceptual user interface for a submarine periscope system*. Paper presented at the 10th Industrial and Organisational Psychology Conference, Perth, Western Australia.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12 (1), 97-136.
- Vandekerckhove, J., Rouder, J.N. & Kruschke, J.K. Editorial: Bayesian methods for advancing psychological science. *Psychonomic Bulletin & Review* 25, 1–4 (2018).
<https://doi.org/10.3758/s13423-018-1443-8>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25, 35-57.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129-133.
- Walker, B. H. (2000). Optical design for visual systems. SPIE Press.
- Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15, 160-168.
- Wierwille, W. W., & Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. *Paper presented at the Proceedings of the Human Factors Society Annual Meeting*, 27, 129-133.

- Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in cognitive sciences*, 7(2), 70-76.
- Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28, 1060-1092
- Wolfe, J. M. (2014). *Approaches to visual search: Feature integration theory and guided search*. In A. C. Nobre & S. Kastner (Eds.), Oxford library of psychology. The Oxford handbook of attention (p. 11–55). Oxford University Press.