Analysis of submarine steering: effects of cognitive and perceptual–motor requirements on the mental workload and performance of helmsmen

Philippe Rauffet, Christine Chauvin, Chiara Nistico, Samantha Judas, Norbert Toumelin

ANALYSIS OF SUBMARINE STEERING: EFFECTS OF COGNITIVE AND PERCEPTUAL-MOTOR REQUIREMENTS ON THE MENTAL WORKLOAD AND PERFORMANCE OF HELMSMEN

Philippe Rauffet, Christine Chauvin, Chiara Nistico, Samantha Judas

{philippe.rauffet, christine.chauvin, chiara.nistico, samantha.judas}@univ-ubs.fr;

Université Bretagne-Sud, Lab-STICC UMR CNRS 6285, Lorient, France

Norbert Toumelin

norbert.toumelin@dcnsgroup.com, DCNS Group, Lorient, France

Abstract: The purpose of the study is to model manual submarine steering and its effects on helmsmen in terms of performance and mental workload. This activity is first formalized according to (a) cognitive requirements, by identifying different levels of control, and (b) perceptual-motor requirements, by analysing the directional compatibility of control-display design. An experiment is then carried out on a simulator designed by a world leader in military naval shipbuilding. This experiment follows a unique scenario including two driving situations with different levels of cognitive requirements (approach and stabilization phases). It is achieved by two groups, each carrying out a perceptual-motor task on a specific steering control-display configuration proposed by the naval shipbuilder (one with a standard numeric display and one with a new visual-spatial representation, both tasks controlled by the same joystick). The findings of this study show that the cognitive requirements during the stabilization phase with high propulsion speed produce increased mental workload, and the perceptual-motor requirements also produce increased mental workload when a direction-of-motion stereotype is violated (upward-forward relationship).

1. INTRODUCTION

Submarine steering is a specific case in activities of complex system control. Controlling the three-dimensional movement of military submarines is carried out without direct visualization, neither of the environment nor of the submersible actuators. Moreover, activities consist in a set of routine procedures and phases carried out by a relatively inexperienced helmsman responding to compass heading and immersion instructions transmitted by the chief of the watch (COW). This relative inexperience is explained by the low

level of decision taken by these operators, who only execute the instructions. The position of helmsman is one of the first appointment of young submariners. Helmsmen are generally recruited with an A-level diploma, and they are assigned to steering tasks just after a short formation lasting 8 weeks, whose 20 hours spent on a simulator.

A few authors have investigated the working activities carried out in a submarine control room. For instance, McLane and Wolfe (1967) focused on display concepts for submarine manual control, and more recently, Stanton and Bessell (2014) proposed a work analysis of the activities involved in returning the submarine to periscope depth. However, these analyses relative to submarine control were mostly descriptive or narrative, and little concerned with performance and the effects experienced by the driver (Verney, 1995). The lack of in-depth studies is explained by the difficulties of accessing the domain context because of the physical constraints and confidentiality issues. There is, however, a startling contrast with the research conducted in the aviation or automotive domains, where numerous studies provide fine-scale modelling of cognitive steering tasks (McRuer et al., 1977; Ranchet, 2011), an analysis of perceptual-motor interactions with a human-machine steering interface (Temme et al., 2003; Jamson et al., 2005), and a regular or periodic, objective or subjective, assessment of the effects of steering on performance and mental workload. This assessment is often carried out in simulation environments with strong ecological validity, and it is based on, inter alia, physiological and ocular measurements or questionnaires such as the NASA Task Load Index (NASA-TLX) to determine mental workload (Engstrom et al., 2005; Recarte et al., 2008, Nilsson et al., 2009; Chen & Epps, 2014).

The purpose of the present article is thus to partially open this "black box" by analysing submarine steering and measuring its effects on the helmsmen's performance and mental workload. Submarine steering can be carried out according to different modes ranging from full automation to complete manual control. This paper deals with the manual mode only, because it is the most "costly" for helmsmen in terms of task load. To this end, through a simulator-based experiment, we present an analysis of helmsmen's cognitive and perceptual-motor tasks in an ecologically valid scenario.

2. ANALYSIS OF MANUAL SUBMARINE STEERING TASKS

In this section, we first present a state-of-the-art review of models related to generic steering tasks and direction-of-motion stereotypes. For both aspects, we transpose the models to the case of submarine steering in order to analyse cognitive and perceptual-motor requirements, namely the way pilots represent the situation of the submarine in its environments (2.1) and the way pilots interact manually with steering controls and displays (2.2). Finally, we present studies about the assessment of steering effects on human performance and mental workload (2.3).

2.1. COGNITIVE REQUIREMENTS: FROM GENERIC STEERING TO SUBMARINE STEERING

Driving and steering activities are part of the class of control activities in a dynamic and uncertain situation. A situation is considered as dynamic when the state and the salience of the objects within the environment change over time (Foulsham et al., 2010), and uncertain because the agents partly control these objects (Hoc, 2001). These situations concern industrial process control, air traffic control, highly automated aircraft piloting, etc. These actions are carried out within and upon systems presenting different levels of complexity. Hoc (1996) identifies five factors of complexity: the field of supervision (the scope of the information and available actions), the more or less immediate type of control (is the effect of a command on the process a direct one or is it necessary to know a particular causal chain to anticipate the effects of the chosen action?), the availability of information, the temporal constraints (does the process to be monitored evolve slowly or rapidly?), and the level of automation. Hence, temporal constraints enable the distinction between the driving activities of slow processes (ships) and the driving activities of rapid processes (aircraft). Moreover, these activities are carried out in situations of high uncertainty, where risks are omnipresent, and operators must divide their time among several tasks; the multiplicity of tasks originates from exogenous sources (e.g. the simultaneously existence of several conflicts in air traffic control) and endogenous sources (different levels of cognitive control).

Generic models about steering levels and proactive and reactive controls
Since the work of Michon (1985), steering tasks have been broken down into several levels. Michon's model, developed to account for car driving, identifies a *strategic* level, a *tactical* level, and an *operational* level. The strategic level involves the overall planning of movement. It includes the goals of the trip, the route selection, and an evaluation of associated costs and risks. The tactical level involves the manoeuvres carried out by the driver (avoiding obstacles, overtaking, lane changing). At the operational level, the drivers ensure the lateral and longitudinal control of their vehicle.

Ranchet (2011) links these different levels to the cognitive functions activated by drivers. At the strategic level, the tasks require an intense cognitive investment and the availability of attention resources. Temporal pressure, however, is comparatively low. The itinerary planning, navigation, or route selection tasks call upon high-level executive competencies such as planning, organization, and complex reasoning. At the tactical level, the tasks involve decision-making regarding the actions that are going to be performed within a few minutes. The task cognitive costs are lower than those at the higher level, but their time requirement is relatively high. This level brings together various vehicle steering actions, such as overtaking, maintaining inter-car distance, adjusting speed, or anticipating situations (bends, traffic lights, etc.). These behaviours require executive competencies, such as planning and task switching, and lower level cognitive competencies, such as information processing speed. Finally, the operational level corresponds to the execution of actions decided at the higher levels. This level is lower in the model hierarchy, and the tasks arising from that level are mostly or in part automated. Their cognitive cost is relatively low, but their time

requirement is very high. This level contains numerous tasks related to guidance, trajectory control, responses to obstacles, or the motor task completion of manoeuvres. These behaviours call for information processing speed and visual and perceptual capacities.

Furthermore, steering activities have led to the design of specific models based on control theory in the domains of car driving (McRuer et al., 1977) and ship navigating (Anzai, 1984; Kok & Stassen, 1980). The authors emphasized that a control task is difficult, in spite of the apparent simplicity of the environment. Anzai (1984) pointed out four sources of difficulties: the dynamic character of the system (inertia in particular), its rate control (due to the rate control, the effect of past actions persists for a long time), the response time lag, and the difficulty of steering straight. Operators must acquire cognitive strategies for overcoming these difficulties and in particular, a causal knowledge that associates a control action with a perceived change of the ship's trajectory. In the framework of cognitive system engineering, Hollnagel (1993) also identified several modes of cognitive control, from the most reactive (determined by the occurrence of external data and characterized by a short time process) to the most proactive or anticipative (relying on the process of internal data and characterized by a longer time). According to Hoc and Amalberti (2007, p.27), "situation mastery and adaptation only marginally relate to a closed-loop reactive model. Rather, they are based on an open-loop anticipative model".

The specific case of manual submarine steering

The distinctive feature of submarine steering is the three-dimensional environment of the vehicle. The steering parameters are compass heading (direction in the horizontal plane) and immersion (vertical position). Adjusting these parameters involves implementing two kinds of forces (Verney, 1995): static forces (submarine buoyancy is achieved through controlled filling of the ballast tanks) and dynamic forces (helm orientation creates lift forces proportional to the square of the speed in relation to the water). Moreover, this movement needs to be controlled in terms of the submarine requirements of security, stealth, and manoeuvrability. The parameters of steering expertise are rate of climb (vertical movement), trim of the vessel (the angle of the submarine relative to the horizon plane), and the ratio between the propulsion speed and immersion depth (the speed-depth domain is a chart specifying the environmental physical constraints on the submarine, and pressure forces in particular). Finally, submarine steering is carried out by helmsmen (who will often be trained submariners with little experience) who execute a set of routine tasks to meet the compass heading and immersion instructions given by the Chief of the Watch (COW) or the officer of the deck or to respond to specific incidents (damage, flooding, fire). Meeting instructions, namely reaching set values, can be broken down into two phases:

- a phase of approach, where the helmsmen seek to go quickly close to the immersion instructions from a different initial immersion state. This phase corresponds to the manoeuvres of rising and diving, where the operators have to control the rate of climb and trim of vessel, so that they do not exceed thresholds above which submarine control could be lost.

- a phase of countersteering and stabilization, where the helmsmen seek to reach the reference values accurately and asymptotically. The operators have to anticipate submarine inertia and dynamics to prevent the submarine exceeding immersion instructions and showing an unstable, oscillatory behaviour.

Application of generic steering models to understand cognitive requirements in manual submarine steering

In this study, following Ranchet's (2011) model, we focus on the tactical control level (control of the position and direction relative to the compass heading and immersion instructions) and operational control level (control of trim, rate of climb, system capabilities, and environmental constraints) in manual submarine steering by helmsmen. The strategic level is not part of the helmsmen's responsibilities, since the objectives and movement planning are determined by the officer of the deck and the COW.

Figure 1 shows these three steering levels that allow a first division of helmsmen's activities in steering situations. Table 1 shows the classification of steering parameters in terms of operational and tactical levels.

Table 1 – Classification of steering parameters for operational and tactical levels

Steering levels	Functions	Steering parameters
Operational	Control of movement	Trim of the vessel
		Rate of climb
	Control of system capacities	Helm / maneuvering space
	Control of system/environment interactions	Speed-depth domain / propulsion speed / immersion minimum
Tactical	Complying with instructions	Immersion depth
		Compass heading

The tactical and operational levels are both continuously brought into play throughout a manoeuvre and require from the helmsmen cognitive, motor, and perceptual capacities that should be seen as highly interactive (Peters & Nilsson, 2007). Furthermore, following Hollnagel's (1993) dichotomy, we consider the approach phases as falling within reactive control (checking the state of the submarine against the instructions) and the stabilization phases as falling within proactive control (because of the response time between the reaction of the submarine and the pilots' action, the latter need to anticipate the behaviour of the vehicle). This analysis of proactive and reactive modes allows a second division of steering situations. Within tactical and operational levels, we identify two phases, shown in Figure 1.
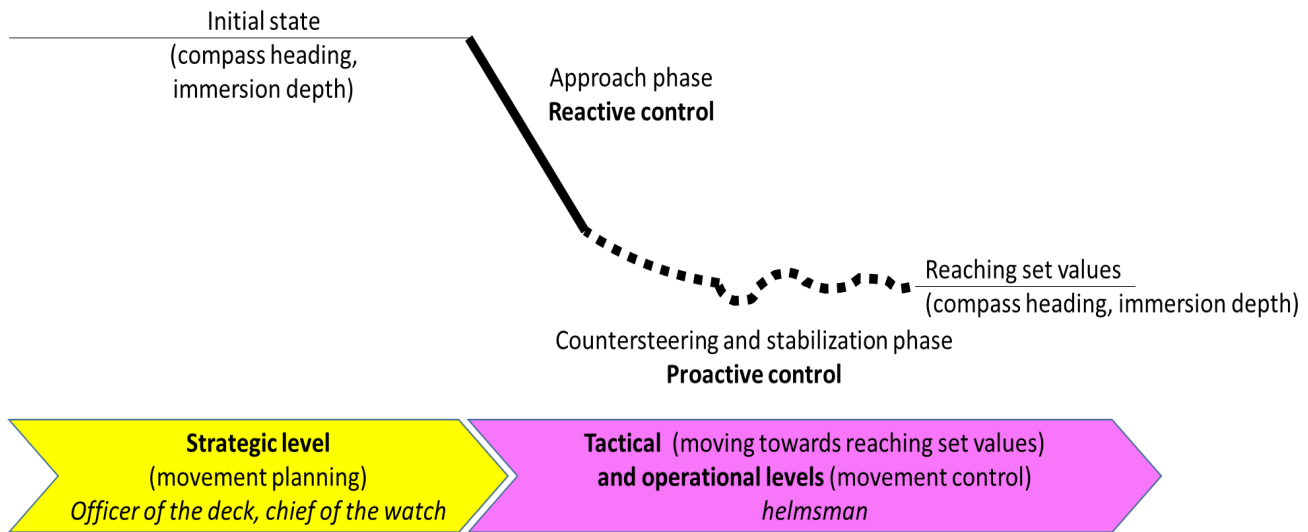
Figure 1 – Typologies of steering situations

2.2. PERCEPTUAL AND MOTOR REQUIREMENTS: CONTROL-DISPLAY COMPATIBILITY

In the later generations of submarines, all the elements needed for helm supervision (control and display) are found in one unit: the steering station. The monitoring and the command of three-dimensional steering parameters (immersion gap, compass heading deviation, trim of the vessel), within tactical and operational phases, involve low-level perceptual-motor tasks with the steering control-display configuration (Verney, 1995, Peters & Nilsson, 2007). These perceptual-motor tasks are often dependent on the activation of a direction-of-motion stereotype, i.e. the expectations of the relevant user population in terms of directional control-display compatibility (Williams, 2004).

The following paragraphs explore these generic stereotype issues that are then used to compare the perceptual-motor tasks resulting from the interactions with two different control-display configurations for submarine steering.

Generic models about direction-of-motion stereotypes

Once the physical requirements for display readability and control motions have been met, the direction-of-motion stereotype or movement compatibility is a preponderant key factor for the successful design of controls and displays (Moray, 1997). The design of systems that do not follow stereotypes reduces performance or safety, especially in emergency situations or under time pressure (Zupanc et al., 2007). An important study evaluating possible types of directional stimulus-response compatibility for systems, in which a control must be moved to bring about the goal-oriented motion of an object, is that of Worringham and Beringer (1998); the issue is also dealt with in the work of Burgess-Limerick et al. (2010) and Chan and Hoffmann (2010). In particular, these authors refer to two common stereotypes that may exist between control and display movements and should guide the design of compatible configurations:

- ***Control-display compatibility (CD):*** the directional relationships in which the control and resulting movement on the display moved in the same direction were the least likely to cause error.

- Visual Field compatibility (VF): The principle suggests that compatible control–response relationships are those in which the direction of movement of the response in the operator's visual field matches the direction of movement of the control in the operator's visual field if the operator is looking at the control. This principle has been demonstrated to predict task performance correctly.

Wickens et al. (2005) also analyse these direction-of-motion stereotypes when visual spatial thinking requires transformations between the different reference frames of the control and the display. In particular, they investigate the case where the control is in the horizontal frame whereas the display is vertically oriented. Contrary to the rather symmetrical lateral axis, where left and right are easy to confuse, the vertical and the fore-aft axes share an asymmetry (due to the gravitational forces, up is different from down, and because of the vision and the locomotion systems, forward is different from backward). As a consequence, the authors notice a strong association between the two marked endpoints of these two asymmetrical axes, especially between "forward" and "upward" directions.

The specific case of two specific submarine steering "control-display" configurations

As part of the development of a submarine, studies were conducted in close collaboration with a shipbuilder. Work focused on the design of a new steering configuration, i.e. a visualization interface coupled with a joystick as a motion control. This new configuration (henceforth called configuration B) could be deployed in future boats, as an alternative to the standard interface currently in use in submarines (henceforth called configuration A). As shown in Table 2, the joystick must be pushed for diving and pulled for rising.

The interface of configuration B aims at replacing the numerical data related to steering parameters by a visual-spatial information synthesis. In particular, it uses certain proposals from Temme et al. (2010) who used two metaphors to build an aircraft cockpit display for the US Naval Air Station. These metaphors were adapted to the submarine case:

- The "starfield" metaphor shows the system's behaviour and its localization in the external world. It is a reference framework that showcases the vertical and horizontal translations as well as the transverse rotations of the submarine.

- The "aircraft" metaphor (in the present case, "submarine") is a 2D image that symbolizes the vessel. Its position is fixed on the screen; the relative position of the starfield metaphor shows the vessel's behaviour and trajectory. Breakdowns (e.g. jammed rudder blades) may also be shown on this metaphor, which accounts for the orientation and movement limits of the system.

Application of stereotype models to the two control-display configurations

Table 2 shows a comparative analysis of the two control-display configurations A and B, where only the interface changes (the joystick stays the same).

Table 2 – Comparison of the two control-display configurations in terms of direction-of-motion stereotype
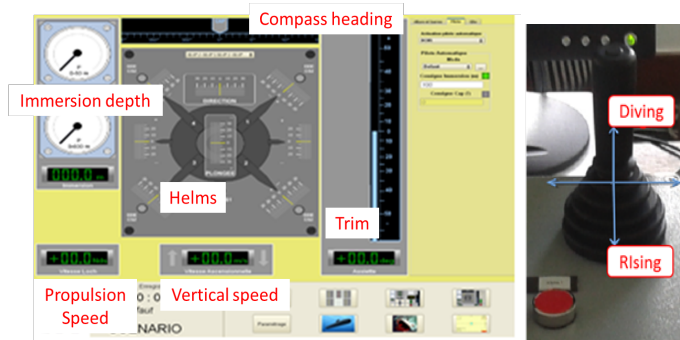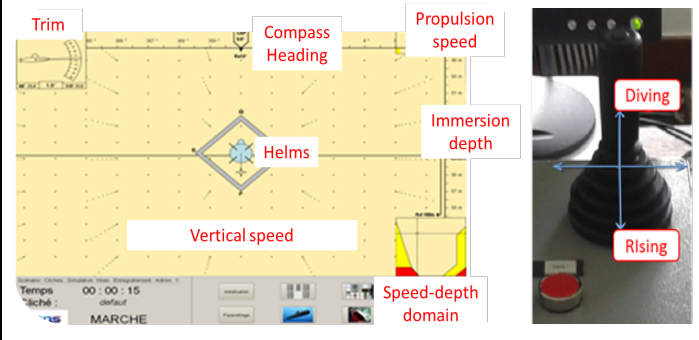
Configuration A	Configuration B
No direction-of-motion stereotype is involved between control and display. The interface is spatially static; only digits change over time for vertical speed and compass heading.	There is a compatibility of control and display for the compass heading monitoring, but there is also a control-display incompatibility for the vertical speed control. The joystick must be pushed forward to dive, implying vertical descending of the diamond-star inside the aircraft metaphor on the display. Hence, the upward/forward stereotype is violated (see Wickens et al., 2005).

The stereotype violation in configuration B was confirmed by the feedback of submariners of two French naval bases, to whom this new configuration was submitted. Several submariners pulled the joystick to increase the submarine immersion, which caused the opposite effect of the desired goal. It is worth remembering that when pushing the joystick, the star moves towards the lower end of the submarine and the latter dives, whereas when pulling the joystick, the star goes towards the high end of the diamond shape and the submarine rises (see Figure 2).
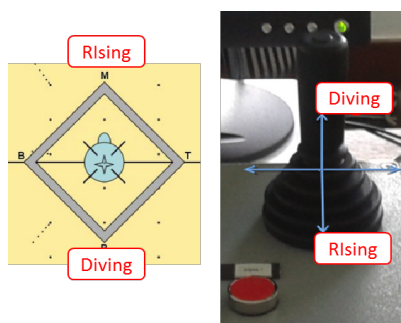
Figure 2 - The disturbing reverse motions in configuration B

This control system thus follows the usual pattern, particularly in the aviation domain: when pilots pull back the stick, the aircraft ascends, and when they push, it descends. The interface, more precisely the diamond-star set-up, may have misled some participants. They may have pulled back the joystick to try and bring the star towards the bottom of the diamond: "*When you try to make the star go up, you pull the joystick and then you reverse the command*". At the opposite, the movements of control and display are parallels for counter clockwise and clockwise turns.

2.3. ANALYSIS OF HELMSMEN'S PERFORMANCE AND MENTAL WORKLOAD

The foregoing literature review regarding generic steering activities enable us to make hypotheses related to the cognitive and perceptual-motor aspects of submarine steering tasks. There are two distinct phases (stabilization and approach) with different cognitive requirements, and the perceptual-motor tasks require compliance with the "upward/forward" direction-of-motion stereotype. To assess the effects of these different aspects of steering tasks, helmsmen's performance and mental workload need to be measured. According to the literature, different indicators can be identified.

Steering performance indicators

Helmsmen need two main skills for manual steering tasks, related to manoeuvre execution speed and directional stability. On the one hand, as for surface vessels, submarine pilots have to face the "response time lag" (Anzai, 1984) of the submarine. They should be able to perform prescribed manoeuvres as quickly as possible (Sutton & Towill, 1986), in compliance with physical limits like trim of vessel, so as to be able to face emergency situations (presence of other ships or obstacles in the immediate environment of the vessel, damages or fire, etc.). On the other hand, helmsmen must "steer straight" and control the vertical and horizontal "turn rate" (Anzai, 1984). They have to control the immersion and compass heading and to optimize the linear route of the submarine, so as to prevent the vessel from directional instability and potential decrease of ship stealth (Koyoma et al., 1978).

Helmsmen's mental workload indicators

Many studies propose subjective and objective measures to assess mental workload. Among the subjective techniques, there are two main multi-dimensional methods: SWAT (Reid & Nygren, 1988) and NASA-TLX (Hart & Staveland, 1988). When the SWAT scale is compared to the NASA-TLX, the TLX scale is generally considered to be the better scale for measuring mental workload (Hill et al., 1992). The NASA-TLX rating scale (TLX stands for task load index) is designed to assess the different factors causing mental workload in a questionnaire that enables a periodic, multidimensional, and subjective evaluation of the load experienced. This questionnaire examines six dimensions of workload assessment: cognitive demand, physical demand, temporal demand, effort, performance, and frustration. Hence, the different dimensions characterize the perception of the demand by the participants and their cognitive activity and effort. At the end of the main task, the participants score each dimension from 0 to 100. These six dimensions are then displayed in pairs, and the participants select the dimension that contributed most to their load. Byers et al. (1989) proposed an alternative to Hart and Staveland's (1988) NASA-TLX, called NASA-rTLX (standing for raw task load index), whereby the weight gradings of the different dimensions are identical instead of being weighted by each subject. The authors showed a strong correlation between the two indicators.

In addition to this subjective assessment, many studies have shown the contribution of objective, physiological measures for estimating mental workload, such as heart rate or skin conductance (Veltman &

Gaillard, 1996, Wilson, 2002). In particular, several authors focused on another physiological indicator, the pupillary diameter, because remote eye tracking sensors are minimally intrusive and do not interfere with operators' activity. Hence, Beatty and Lucero-Wagoner (2000) showed that an increase of the mean value of pupillary diameter results from the cognitive effort increase needed to face a more difficult task. Cegarra and Chevalier (2008) argued that pupillary response can be considered a very sensitive indicator that can complement and be combined with the NASA-TLX technique. Marshall (2007), Dehais et al. (2008), Recarte et al. (2008), De Greef et al. (2009) also identified this relationship in real or simulated activities, under various driving or steering tasks such as piloting an airplane, driving a car, or operating a naval warship.

The next paragraphs present an experiment designed to verify the modelling of the specific case of manual submarine steering proposed in sections 2.1 and 2.2 and to determine precisely the effect of this activity (both cognitive and perceptual-motor tasks) on the performance and the mental workload of helmsmen with the indicators presented in section 2.3.

3. MATERIAL AND METHOD

A first experiment designed to measure the effect of steering tasks upon mental workload and performance had already been conducted with 20-odd submariner helmsmen in two French naval bases. The sample, however, was highly heterogeneous, whether in terms of length of service in the Navy (ranging from 18 months to 10 years), campaigns at sea, or job profiles (certain participants had received helmsman training only, and others had added COW or officer of the deck training). Moreover, configuration A (see 2.2) was familiar to all these submariners, which consequently skewed the configuration comparison. Additionally, a helmsman's job is to execute simple instructions; it can be entrusted to inexperienced submariners.

Hence, we decided to reproduce the experiment with a novice homogeneous population having a solid scientific knowledge level. The population knew neither of the two steering interfaces (hence the bias of the first experiment was avoided) but was knowledgeable enough to comprehend and control the dynamic behaviour of the submarine; the group of participants thus closely resembled the average submarine helmsman profile.

The experiment protocol is detailed in the following sections.

3.1. PARTICIPANTS

The participants were recruited from undergraduate and doctoral students in an engineering school in Brittany, France. Twenty students participated in the study (average age = 23.80; standard deviation = 3.105); they were randomly distributed into two equal groups corresponding to the two experimental conditions: one group carried out the simulation with configuration A, the other with configuration B (see

2.2). Furthermore, all participants worked with the same experimental scenario (see 3.2). Group homogeneity was controlled through two variables: video game playing and visual correction. Both groups had the same number of participants who did not play video games (N=3), the same number of participants who played video games at least once a week (N=3), and the same number of participants wearing contact lenses or glasses (N=4).

3.2. SIMULATION SET-UP AND EXPERIMENTAL SITUATION

The simulator

The simulation set-up was designed to reproduce an ecological situation:

- *at the operating level*, the participants used a joystick; this device corresponds to what will replace the current supporting-bracket system in the future submarines;

- *at the visualisation level*, the interface of configuration A reproduced the current interface in submarines but on a smaller scale, and the interface of configuration B was an alternative tested by the shipbuilder for the future submarines;

- *finally, at the behaviour level*, the software part of the simulator reproduced a submarine's hydrodynamic model for both standard manoeuvres and degraded situations (technical breakdown, damage, etc.). In accordance with the instructions issued through the commands, the simulation software computed the submarine behaviour in real time and displayed the information as it normally would appear on the submarine steering station.

The scenario

For this study, a scenario with four steering manoeuvres was developed and validated in association with a submarine captain (see Figure 3). The values of speed, immersion, and trim limits of the scenario were chosen with this expert; the French Navigation School of Submariners also validated their ecological coherence. The four manoeuvres result from the repetition of two sequences; each sequence included one submersion manoeuvre carried out at fast propulsion speed (16-20 knots) and one surfacing manoeuvre carried out at moderate speed (8-10 knots).

These manoeuvres reflected the pilots' response to instructions regarding immersion depth changes (+/- 50m on a relative basis); their difficulty depended upon the propulsion speed and the steering phase (i.e. steering time requirements and reactive approach or proactive stabilization (see 2.1)), manipulated by the experimenter. Participants had to reach the ordered immersion depth (within +/- 7m of receiving the instruction) and stabilize the submarine for 10s so that the manoeuvre was evaluated as completed.

There was no time limit, but each participant was given a maximum of four attempts per manoeuvre to reach immersion depth. If all four attempts were used, the experimenter ended the manoeuvre. This threshold of four attempts was chosen for a practical reason, namely to prevent underperforming participants from being

subjected to never-ending sessions. Moreover, in agreement with the submarine captain helping to design the scenario, it was considered as a good indicator of the operators' (in)ability to achieve the requested task. To change immersion depth, pilots act on the trim of the vessel only, hence on the submarine's rate of climb.
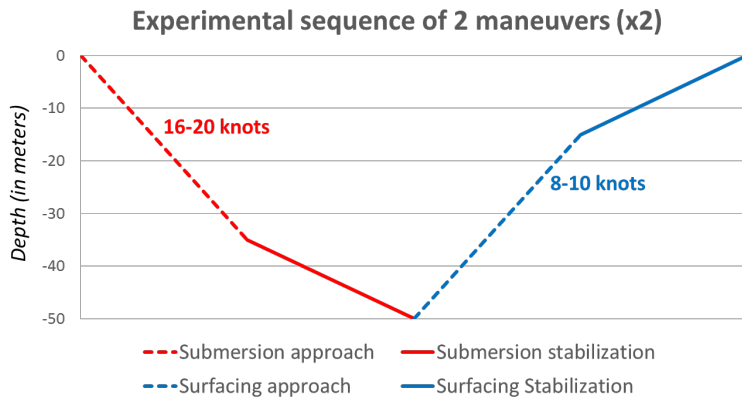

Figure 3 – Experimental scenario

3.3. DATA COLLECTION AND PROCESSING

Two independent variables were used: performance and mental workload.

Measure of performance

During each manoeuvre, the participants' performance was evaluated. Three parameters, related to the manoeuvre execution speed and submarine route stability (see section 2.3), were defined with the expert who validated the scenario:

- ***ImmOK:*** reaching depth and stabilizing according to immersion instructions (after a maximum of four attempts). ImmOK takes the value 1 if the manoeuvre is successful and 0 otherwise.

- ***NbAtt:*** number of attempts before reaching immersion depth; beyond the limit of a maximum of four attempts, the manoeuvre was interrupted and this failure was coded as a fifth attempt.

- ***SpeedExec:*** speed of execution of each phase or each manoeuvre.

Measure of mental workload

Mental workload was evaluated with the aid of subjective and objective measures, following the literature techniques detailed in section 2.3.

Using the NASA-rtlx to measure mental workload

After each manoeuvre, participants answered the French, pen-and-paper version of the NASA-rtlx questionnaire (cf. 2.3). The questionnaire retained its six dimensions. For each of these, participants answered on a Likert-type scale ranging from weak to high levels. Each dimension was explained verbally.

Using eye-tracking data to estimate mental workload

For this study, we used a binocular eye-tracker with a 60Hz recording frequency allowing precision readings between 0.5° and 1.0°. The recorded data were then processed using Seeing Machines software:

 - *faceLAB™* 5 enables the continuous measure of pupil diameter, head posture, and gaze direction.

 - *EyeWorks™* provides the link between the ocular data and the experimental environment. It provides screenshot video capture and records gaze direction and data breakdown according to the phases of the scenario or areas of interest.

The eye-tracker was placed at the bottom of the screen and behind the control box (see Figure 4). To avoid ocular behaviour variations (pupil diameter, blinks, etc.) resulting from changes in ambient light conditions, the luminosity of the test venue was kept constant.

Figure 4 – Experimental set-up

The raw data were cleansed and processed with MATLAB software (see Table 3).

Consistent with the literature, physiological ocular data (pupil diameter) are considered an indicator of mental workload. Only pupil diameters between 2 and 8 mm were retained (which corresponds to the maximal dilation range of the human eye); similarly, fixation time was taken into account from 15 hundredths of a second (0.15s) only. Pupil diameter was also subject to centring-reducing operations so as to eliminate inter-subject differences related to average diameter and dilation variations.

Synthesis of the selected indicators

Dividing each manoeuvre into approach and stabilization phases was performed by coding the starting times of these phases from the video captures. More precisely, the start times of approach phases were coded according to the input of immersion instructions in the simulator and corresponding to the running simulator. The end times of approach phases and the start times of stabilization phases corresponded to the first entrance of the submarine in an interval of +/- 7m around the immersion instructions. Finally, the end times of stabilization were coded by establishing when the experimenters paused the simulation, which indicated either the success of stabilization for 10s, or a failure after four attempts.

This division enabled the calculation of indicators for each phase; these were the dependent variables (DVs) under investigation in this study.

Table3 – Synthesis of the dependent variables

Variable	Raw data	Filtered and standardized data	Calculation
ImmOK *(for each phase)*	Immersion success or failure (4 attempts max)		$ImmOK = \begin{bmatrix} 0\ if\ failure\ after\ 4\ attempts \\ 1\ if\ success \end{bmatrix}$
NbAtt *(for each stabilization phase)*	Number of attempts		$NbAtt = \sum_{attempt\,\in\,stabilization\ phase} attempt$
SpeedExec *(for each phase or each maneuver)*	Duration of execution		$SpeedExec = \dfrac{Distance\ to\ cover}{Duration\ of\ execution}$
NASA − rtlx *(for each maneuver)*	Likert-type scale for each dimension	Overall indicator calculated from the mean of all dimensions	$NASA - rtlx = \begin{bmatrix} experienced\ physical\ demand \\ experienced\ mental\ demand \\ experienced\ temporal\ demand \\ experienced\ effort \\ experienced\ performance \\ experienced\ frustration \\ overall\ indicator\ of\ experienced\ load \end{bmatrix}$
PD *(for each phase)*	Raw pupil diameter	Cleaned and z-normalized pupil Diameter DPz (Diameter between 2 and 8 mm, then z-normalizing operations)	Mean pupil diameter for each phase $PD = \sum_{DPz\,\in\,phase}^{1..n} \dfrac{PDz}{n}$

Statistical processing

Analysis of the distribution of the dependent variables and of the control variables then helped determine the nature of the applicable statistical tests.

The conditions of normality, homogeneity, and sphericity of variances were met by only the mean of z-normalized pupil diameter and certain dimensions of NASA-rtlx. Repeated measures univariate ANOVA and a *t*-test were then applied using STATISTICA software.

Regarding the other variables, the nonparametric Mann-Whitney U test was used to compare the means of the two groups.

3.3. EXPERIMENTAL PROTOCOL

The experiment was conducted in the LOUSTIC laboratory, a multidisciplinary research platform for the uses of information and communication technology. Participants took the test individually. The test lasted about 2 hours and was divided into six phases:

1. greeting participants: completing the profile questionnaire;

2. explaining the steering principles: a slide presentation was used to explain (verbally) the steering tasks and the interface;

3. practicing: participants carried out three manoeuvres to familiarize themselves with the steering activities, the control box, the interface, and the important navigation parameters to monitor. This short training session lasted about five minutes and corresponded to the achievement of one diving manoeuvre for each participant. Moreover the training manoeuvres were carried at a very low propulsion speed (around 4

knots) in comparison with the scenario speeds. The training was designed to help participants understand submarine dynamic situations and build a mental model to understand vessel inertia and physical limits. However, the slow speed and the short length of this practicing session were chosen to prevent participants from accustoming to the scenario conditions, so is to say to avoid a learning effect that could have an impact on the steering performance during the experiment.

4. eye-tracking parameter and calibration setting: Eye-Works Record software was launched to record the participants' eye movements after faceLAB software had calibrated head, eyes, and test environment;

5. carrying-out the steering task: participants carried out the manoeuvres and completed the NASA-rtlx questionnaire at the end of each manoeuvre;

6. debriefing and thanking: participants were questioned regarding the set of manoeuvres, the interface, and the simulator in order to obtain their comments, then the test purpose was revealed. To ensure the inter-subject independence of the collected data, participants were asked not to reveal test contents to those around them.

3.5. EXPERIMENTAL DESIGN

The protocols of only 16 out of 20 participants were analysed because of a problem of pupil diameter data acquisition for certain manoeuvres. In the end, the configuration A group included 8 participants and the configuration B group included 8 individuals; the control-display configuration thus constituted an inter-subject variable. All were subject to the same scenario (sequence, type of manoeuvre, and type of phase were intra-subject variables).

Our experimental design was thus a mixed factorial design, written as in *Participant8<ControlDisplayConfiguration2>*Sequence2*Manœuvre2*Phase2*. The 16 participants were therefore divided into two independent groups; each group experienced a single control display configuration only. The same scenario was submitted to all the participants, who were assigned to two repeated sequences composed of two distinct manoeuvres, a rapid submersion and a moderate speed surfacing. In each manoeuvre, participants achieved two phases, namely one approach and one stabilization. The tables presented below show these independent variables (related to cognitive and perceptual-motor requirements), the dependent variables (related to performance and mental workload), and the control variables (operators' behaviour).

Table 4 – Independent and dependent variables

Dependent variables	Indicator of	Ranges
ImmOK : Immersion success	Performance	Binary (0 or 1)
NbAtt : number of attempts	Performance	Integer (from 1 to 4)
SpeedExec : Execution speed for the phase or the maneuver in m.s^{-1}	Performance	Positive real numbers (min: 0.007; max: 1.023; mean: 0.214; std: 0.117)
NASA-rtlx : dimensions of experienced workload	Mental workload	Integer, for each dimensions and for overall workload (from 0 to 100, with a step of 5)
PDCR : pupil diameter (z-normalized)	Mental workload	Real numbers (min:-0.658; max: 1.481; mean: 0.113; std: 0.331)

Independent variables	Groups	Categories
Control-Display Configuration	2 Independent	Configuration A Configuration B
Sequence	2 Paired	Sequence 1, Sequence 2
Maneuver	2 Paired	Submersion at rapid speed, Surfacing at moderate speed
Phase	2 Paired	Approach, Stabilization

3.6. HYPOTHESES

In terms of the distinction between cognitive tasks within driving situations (situation assessment difficulty), modelled in 2.1, and of the distinction between perceptual-motor tasks with different control-display configurations (stereotype compliance or violation), presented in 2.2, we posit the following three hypotheses related to the effects of the steering activity upon the helmsmen's performance and mental workload.

H1 – Effect of the perceptual-motor requirements upon performance

The control-display configuration has an effect upon the operators' performance. We expect configuration B to decrease participants' performance significantly, due to the violation of the "forward-upward" stereotype and the resulting increased difficulty of perceptual-motor tasks.

H2 – Effect of the perceptual-motor requirements upon mental workload

The control-display configuration has an effect upon the operators' mental workload. The mental workload experienced by operators for steering tasks should be higher for participants using configuration B due to the control-display incompatibility. Hypothesis H2 is broken down into two operational hypotheses:

H2.1 – The interface type has an effect upon mental workload as measured by the NASA-TLX. We expect the NASA-TLX score to be higher for participants using configuration B.

H2.2 – The interface type has an effect upon mental workload as measured from pupil dilation. The pupil diameter of participants using configuration B should be higher than that of participants using configuration A.

H3 – Effect of the cognitive requirements upon mental workload

The cognitive requirements have an effect upon the operators' mental workload. Hypothesis H3 is also broken down into two parts:

H3.1 – Increasing the manoeuvre difficulty (increasing the propulsion speed) will increase mental workload (pupil diameter increase).

H3.2 – The stabilization phase, which requires sensorimotor capacities and proactive monitoring, will involve higher mental workload than the approach phase. We thus expect a significant increase of pupil diameter when measured during the second part of the manoeuvre.

4. RESULTS

4.1. EFFECT OF THE PERCEPTUAL-MOTOR REQUIREMENTS UPON PERFORMANCE

Reaching immersion depth (ImmOK)

To evaluate the participants' performance, reaching the immersion depth required was recorded for each manoeuvre in terms of failure or success for all four manoeuvres.

Analyses showed no significant effect of the interface upon reaching immersion depth. However, for the most delicate manoeuvres (the first and the last), a difference appears with descriptive statistics. Only 60% of the users of configuration B succeeded in finishing the first manoeuvre as against 100% of configuration A. As far as the third manoeuvre is concerned, only 30% of configuration B users succeeded as against 80% of configuration A users.

Number of attempts to reach immersion depth (NbAtt)

To obtain a finer-grained analysis of the participants' performance, the number of attempts to reach the required depth was monitored and recorded. The scenario called for a maximum number of four attempts to reach immersion depth;

The analysis of the number of attempts necessary to reach immersion depth for each manoeuvre showed a significant effect of the control-display configuration upon the number of attempts necessary to complete the first manoeuvre ($U = 22$; $p < 0.05$) and the third manoeuvre ($U = 12$; $p < 0.05$), corresponding to submersion manoeuvres with rapid speed.

Specifically, the average number of attempts to reach the required immersion depth was lower for configuration A users (manoeuvre 1: mean = 2.00; standard deviation = 0.94 – manoeuvre 3: mean = 2.60; standard deviation = 1.50) than for configuration B users (manoeuvre 1: mean = 3.50; standard deviation = 1.50 - manoeuvre 3: mean = 3.00; standard deviation = 1.63) (see Table 5).

Table 5 – Effect of the control-display on the number of attempts

Control-Display Type		Number of attempts to reach immersion depth M1	Number of attempts to reach immersion depth M2	Number of attempts to reach immersion depth M3	Number of attempts to reach immersion depth M4
Configuration	Mean	2.00	1.00	2.90	2.60

A	Standard deviation	.943	.000	1.449	1.506
	N	10	10	10	10
Configuration B	Mean	3.50	1.20	4.30	3.00
	Standard deviation	1.509	.422	1.337	1.633
	N	10	10	10	10
Total	Mean	2.75	1.10	3.60	2.80
	Standard deviation	1.446	.308	1.536	1.542
	N	20	20	20	20

This interaction between the control-display and the manoeuvre difficulty - due to the propulsion speed - is shown in Figure 5 (vertical bars denote 0.95 confidence intervals).
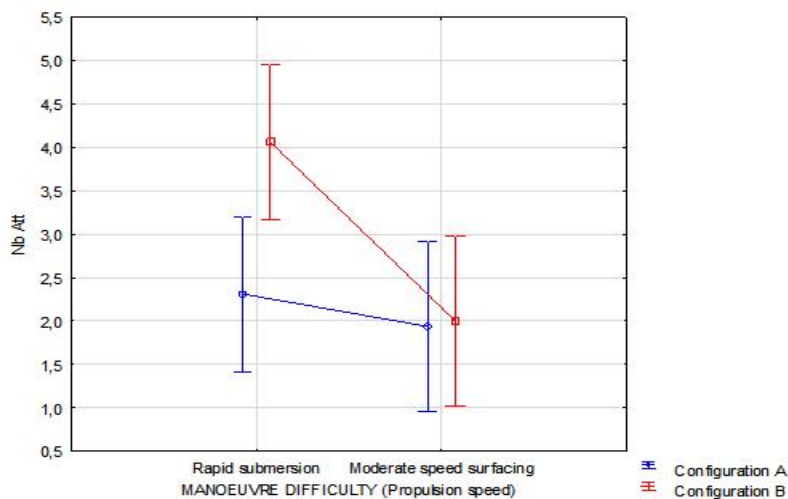

Figure 5 – Effects of the control-display and the propulsion speed difficulty on the number of attempts

Execution speed to reach immersion depth (SpeedExec)
To evaluate performance, the mean speed to complete each manoeuvre was also computed (vertical distance to be covered divided by manoeuvre time). This speed, related to the number of attempts necessary to reach the required immersion depth, can be seen as an indicator of effectiveness. For this variable, analyses did not show any significant effects of the interface.

Effect of sequence on number of attempts and speed to reach immersion depth
The analysis of the last two variables (execution speed and number of attempts to complete manoeuvres) also shows an unexpected effect, namely an effect of sequence (see Table 6).

Table 6 – Effects of sequence upon performance indicators

Significant effects of sequence upon	U	p	Group means
Execution speed (approach, N=64)	U=90	$p=1,5*10^{-8}$	Sequence 1 : 0,15 m/s Sequence 2 : 0,43 m/s
Execution speed (stabilization phase, N=64)	U=343	p=0,024	Sequence 1 : 0,06 m/s Sequence 2 : 0,08 m/s
Number of attempts (stabilization phase, N=64)	U=273	p=0,001	Sequence 1 : 1,93 attempts Sequence 2 : 3,21 attempts

The results shown in Table 6 indicate that:

- participants executed the approach phase three times faster in the second sequence than in the first sequence;

- participants were significantly slower (-30%) to stabilize the submarine in the second sequence than in the first one;

- participants made significantly more errors in the second sequence than in the first one (+66%).

This idiosyncratic behaviour, not controlled by the experimenter, could constitute bias as participants introduced a new parameter of difficulty in addition to the induced independent variables (manoeuvre and phase). Subsequently, the experiment was examined overall but also with a possible focus on the first sequence (without taking into account the repetition and acceleration of manoeuvres induced by the participants' behaviour).

4.2. EFFECT OF COGNITIVE AND PERCEPTUAL-MOTOR REQUIREMENTS UPON MENTAL WORKLOAD

NASA-rtlx

The dimensions investigated in the mental workload questionnaire were mental requirements, physical requirements, time requirement, effort, performance, and frustration (variance homogeneity was controlled and accepted for all dimensions of the questionnaire with the value $\alpha =.05$).

For a macroscopic analysis, the dimension-based means of all manoeuvres were examined (the means normality was controlled and accepted, and variance homogeneity was controlled and accepted for all means of all four manoeuvres with the value $\alpha =.05$). Results show a significant effect of the interface for the mean of the *Effort* dimension ($t (18) = 2.45$; $p < 0.05$). Configuration B users (mean = 70.00; standard deviation = 9.57) report a perceived effort higher than that of configuration A users (mean = 58.75; standard deviation = 10.86).

More detailed analyses were performed to check whether there were differences according to the complexity level of manoeuvres.

Distribution normality was checked, then parametric and nonparametric tests were carried out (normality was respected for all questionnaire dimensions except for the *Effort* dimension in manoeuvre 1).

The score statistical analysis showed a significant effect of the interface for the *Effort* dimension only in the first manoeuvre (Mann-Whitney U Test: U = 1.98; p < 0.05) and the second manoeuvre (t test: t (18) = -2.22; p < 0.05).

The mean values of the *Effort* dimension of the first two manoeuvres are significantly higher for configuration B users (manoeuvre 1 mean = 71.00; standard deviation = 11.25, manoeuvre 2 mean = 68.00; standard deviation = 16.02) than for configuration A users (manoeuvre 1 mean = 55.00; standard deviation = 18.86, manoeuvre 2 mean = 52.50; standard deviation = 15.14). In other words, configuration B brought a higher perceived effort than configuration A.

Assessment of pupil diameter

The analysis of the "mean of z-normalized pupil diameter" variable in terms of the variables interface, sequence, manoeuvre and phase in the entire scenario (N=128) is shown in Table 7.

Table 7 – Effects of the independent variables on pupil diameter

Significant effects of pupil diameter upon	F	p	Power	Group means
Control-Display type	F(1,14)=10,13	0,007	0,84	Configuration A : -0,01
				Configuration B : 0,27
Maneuver	F(1,14)=8,79	0,010	0,79	Rapid : 0,16
				Moderate : 0,01
Sequence*Phase	F(1,14)=22,58	0,0003	0,99	Seq1Appr : -0,05
				Seq1Stab : 0,20
				Seq2 Appr : 0,13
				Seq 2 Stab : 0,06

The analysis highlighted three main observations:

- The pupil diameter mean is significantly higher with configuration B than with configuration A.

- The pupil diameter mean is significantly higher in submersion manoeuvres carried out at fast propulsion speed than in surfacing manoeuvres carried out at moderate propulsion speed.

- There is a significant difference between the mean values of pupil diameter in terms of the "sequence*phase" interaction. The Bonferroni post-hoc test shows that there is a significant difference only between the approach phase and the stabilization phase of the first sequence (p=0.0007).

Because of the change in the helmsmen's operational behaviour when repeating the sequence, we also reproduced the analysis shown above but examining the first sequence only (N=64). We found similar significant effects showing the influence of the factors control-display configuration, manoeuvre, and phase upon the "mean of z-normalized pupil diameter" (summarized and shown in Table 8).

Table 8 – Effects of the independent variables on pupil diameter (first sequence)

Significant effects of pupil diameter upon	F	p	Power	Group means
Control-Display type	F(1,14)=9,72	0,007	0,83	Configuration AC : -0,18
				Configuration BC : 0,33
Maneuvrer	F(1,14)=9,30	0,009	0,81	Rapid : 0,17
				Moderate : -0,02
Phase	F(1,14)=9,33	0,009	0,81	Appr : -0,05
				Stab : 0,20

Figure 6 illustrates these combined effects on pupil (vertical bars denote 0.95 confidence intervals).
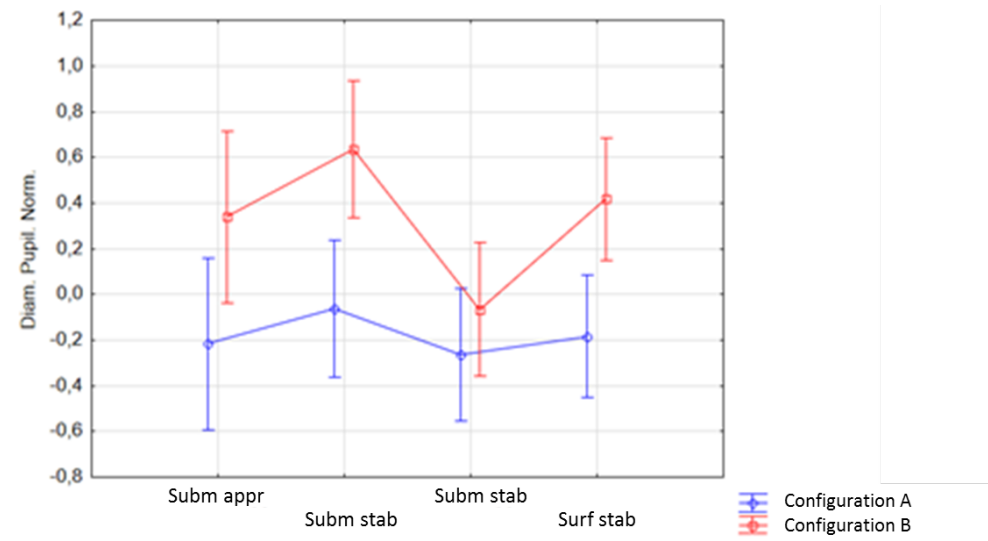

Figure 6 – Effects of the cognitive requirements and the control display in the first sequence

5. DISCUSSION

These findings confirm all the hypotheses posited regarding the effects of the cognitive requirements (situation assessment difficulty) and the perceptual-motor requirements (stereotype compliance or violation) upon performance and mental workload: those effects do exist and are significant.

Consistent with hypothesis H1, there is a significant effect of the perceptual-motor requirements upon performance, due to the violation of a direction-of-motion stereotype. Participants using configuration B had to make more attempts to reach set values, particularly for the more difficult manoeuvres (manoeuvres 1 and 3).

Consistent with hypothesis H2, there is also a significant effect of the perceptual-motor requirements upon mental workload. The participants using configuration B experienced higher mental workload concerning the *Effort* dimension measured by NASA-rtlx. Similarly, the participants' average pupil diameter was

significantly higher than that of the other participants using configuration A, regardless of the manoeuvre or the phase being examined (approach or stabilization phase).

Moreover, consistent with the literature and hypothesis H3, pupil diameter increased with the cognitive requirements. This is attributable to:
- the increased time pressure: increased propulsion speed makes manoeuvring more difficult and thus requires greater control from the helmsmen;
- the need for proactive control: the countersteering and stabilization phase is based upon understanding and anticipation of the dynamic behaviour of the submarine in her environment. The helmsmen thus execute a more complex cognitive process than in the approach phase: not only do they have to correct a trim of the vessel that is too large or monitor the distance from the set values, they also have to find the necessary information to assess the future movement and position of their vehicle.

The statistical results regarding the performance and mental workload experienced with configuration B therefore confirm the comments and the observed behaviours of submariners (see section 2.2). The subjective assessment of these professionals matches completely with the objective measures carried out in our experiment. The joystick moves, opposite to the diamond-star moves on the display of configuration B, increase errors and mental workload. The violation of the upward/forward stereotype defined by Wickens (2005) is consequently demonstrated by these results.

6. CONCLUSION

This experiment has brought useful information on theoretical, methodological, and practical aspects. It has also limitations that should be emphasised. It opens, finally, the way for future research related to the possible automation of the piloting activity that should be presented.

From a theoretical perspective, the experiment validates the steering model we proposed, since it enables a clear distinction to be made between the approach and stabilization phases of submarines, whether achieved at high or medium propulsion speed: the stabilization phase is costlier from a cognitive point of view.
From a methodological perspective, it demonstrates the benefits of combining several measures to enable interface comparison and evaluation. The performance measures and two measures of mental workload (NASA-rtlx and pupil diameter) provided convergent and complementary input for discussion. Following Cegarra and Chevalier (2008), we observe that the pupil diameter measure is more fine-tuned than that of NASA-rtlx (as it is carried out online, it distinguishes the different phases of a manoeuvre) and that NASA-rtlx helps to account for the load participants perceived. In this case, the issue is that of effort, namely the effort that is necessary to go against what appears to be the right action.

From a practical point of view, the experiment has also shown that it is necessary to rethink the control-display configuration B design, so as to comply with the direction-of-motion stereotypes, to simplify the perceptual-motor tasks and to avoid cognitive overload and driving underperformance. That confirms the feedback of professional helmsmen of the interface of configuration B, which mainly pointed out this issue of control-display compatibility (see paragraph 2.2).

However, this study presents some limitations that should be taken into account. It has been conducted in an experimental situation with participants who were not helmsmen. Furthermore, the training period was very short. Although the situation and the scenario were designed in order to be as ecological as possible, the study must be considered as a first approach of the activity of submarine helmsmen. Ideally, this approach should be complemented and strengthened through an analysis of helmsmen's activity in a naturalistic setting. In addition, the study do not consider the difference of interfaces in term of visual design, and it only focus on the control-display compatibility that is the main issue underlined by professional helmsmen. It would be interesting to carry out another experiment, after having removed the stereotype violation on configuration B, in order to investigate further both interfaces and to analyse how the different ways to display information could have an impact on helmsmen performance and workload.

Finally, the experiment results open the discussion on the possible ways to support the helmsmen activity. Future research could deal with the design of ecological interface and/or with new possible forms of automation. On board a submarine, steering may be performed, nowadays, in three different modes, manual (like in the experiment we carried out), automatic (the machine controls immersion depth and heading) or semi-automatic (the machine controls either the immersion depth or heading). The manual mode is used in difficult situations (restricted water or heavy traffic for example). It generates, as seen in this study, a mental workload that could be regulated if the helmsman were assisted in the stabilization phase, with a function guiding the action as a co-pilot could do.

7. REFERENCES

1. Anzai, Y. (1984). *Cognitive Control of Real-Time Event-Driven Systems*. Cognitive Science, 8(3), 221-254.
2. Beatty, J., & Lucero-Wagoner, B. (2000). *The pupillary system*. Handbook of psychophysiology, 2, 142-162.
3. Burgess-Limerick, R., Krupenia, V., Wallis, G., Pratim-Bannerjee, A., & Steiner, L. (2010). *Directional control-response relationships for mining equipment.* Ergonomics, 53(6), 748-757.
4. Byers, J.C., Bittner, A.C. Jr., & Hill, S.G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), Advances in Industrial Ergonomics and Safety (pp. 481-485). London: Taylor & Francis.
5. Cegarra, J., Chevalier, A. (2008*). The Use of Tholos Software for Combining Measures of Mental Workload: Toward Theoretical and Methodological Improvements.* Behavior Research Methods 40 (4): 988–1000.
6. Chan, A.H.S., Hoffmann E.R., (2012). *Movement compatibility for configurations of displays located in three cardinal orientations and ipsilateral, contralateral and overhead controls.* Applied Ergonomics 43 (2012) 128-140
7. Chen, S., Epps, J. (2014). *Using Task-Induced Pupil Diameter and Blink Rate to Infer Cognitive Load.* Human-Computer Interaction. 29(4).

8. Dehais, F., Causse, M., & Pastor, J. (2008, March). *Embedded eye tracker in a real aircraft: new perspectives on pilot/aircraft interaction monitoring.* In Proceedings from the 3rd International Conference on Research in Air Transportation. Fairfax, USA

9. de Greef, T., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). *Eye movement as indicators of mental workload to trigger adaptive automation* (pp. 219-228). Springer Berlin Heidelberg.

10. Engstrom, J., Johansson, E., & Ostlund J. (2005). *Effects of visual and cognitive load in real and simulated motorway driving.* Transportation Research, 8, 97–120.

11. Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). *Gaze allocation in a dynamic situation: Effects of social status and speaking.*Cognition, 117(3), 319-331.

12. Hart S., Staveland, L. (1988). *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research.* Advances in psychology, Elsevier.

13. Hoc, J.M. (1996). *Supervision et Contrôle de Processus: La Cognition En Situation Dynamique.* PUG. Presses universitaires de Grenoble.

14. Hoc, J. M. (2001). *Towards a cognitive approach to human–machine cooperation in dynamic situations.* International journal of human-computer studies, 54(4), 509-540.

15. Hoc, J. M., & Amalberti, R. (2007). *Cognitive control dynamics for reaching a satisficing performance in complex dynamic situations.* Journal of cognitive engineering and decision making, 1(1), 22-55.

16. Hollnagel, E. (1993). *Human reliability analysis: Context and control.* London: Academic Press.

17. Jamson, A. H., & Merat, N. (2005). *Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving.* Transportation Research Part F: Traffic Psychology and Behaviour, 8(2), 79-96.

18. Kok, J. J., & Stassen, H. G. (1980). Human operator control of slowly responding systems: Supervisory control. *J. Cybernet. Info. Sci*, *3*, 123-174.

19. Koyama, T., Kose, K., & Hasegawa, K. (1978). *A Study on the Instability Criterion of the Manual Steering of Ships.* Naval architecture and ocean engineering, 16, 15-24.

20. Marshall, S. P. (2007). *Identifying cognitive state from eye metrics.* Aviation, space, and environmental medicine, 78(Supplement 1), B165-B175

21. McLane, R. C., & Wolf, J. D. (1967). *Symbolic and pictorial displays for submarine control.* Human Factors in Electronics, IEEE Transactions on, (2), 148-158.

22. McRuer, D. T., Allen, R. W., Weir, D. H., & Klein, R. H. (1977). New results in driver steering control models. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *19*(4), 381-397.

23. Michon, J. A. (1985). *A critical view of driver behavior models: What do we know, what should we do?* In L. A. Evans & R. Schwing (Eds.), Proceedings of the International Symposium on Driver Behavior and Traffic Safety. General Motors Research Laboratories (pp. 485-520). New-York: Plenum Press.

24. Moray, N., (1997). *Human factors in Process control.* In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. John Wiley & Sons, Inc, pp.1944-1971.

25. Nilsson, R., Gärling, T., & Lützhöft, M. (2009). *An experimental simulation study of advanced decision support system for ship navigation.* Transportation research part F: traffic psychology and behaviour, 12(3), 188-197.

26. Peters, B., & Nilsson, L. (2007). Modelling the driver in control. In C. Cacciabue (Ed.) *Modelling driver behaviour in automotive environments* (pp. 85-104). Springer London.

27. Ranchet, M. (2011). Effet de la maladie de Parkinson sur la conduite automobile – Implication des fonctions executives. Thèse de doctorat, Université de Lyon 2.

28. Recarte, M. A., Perez, E., Conchillo, A., & Nunes L. M. (2008). *Mental workload and visual impairment: differences between pupil, blink and subjective rating.* The Spanish Journal of Psychology, 11, 374–385.

29. Stanton, N. A., & Bessell, K. (2014*). How a submarine returns to periscope depth: Analysing complex socio-technical systems using Cognitive Work Analysis.* Applied ergonomics, 45(1), 110-125.

30. Sutton, R., & Towill, D. R. (1986). *The Helmsman as a Man–Machine Element.Journal of Navigation*, 39(01), 49-65.

31. Temme, L.A., Still, D.L., Kolen, J., (2003), *OZ: A Human-Centered Computing Cockpit Display.* 45th Annual Conference of the IMTA, Pensacola, Florida

32. Veltman, J. A., & Gaillard, A. W. K. (1996). *Physiological indices of workload in a simulated flight task.* Biological psychology, 42(3), 323-342.

33. Verney, J. (1995). Pilotage intégré pour sous-marins. *Navigation*, *43*(171), 372-387.

34. Wickens, C. D., Vincow, M., & Yeh, M. (2005). *Design applications of visual spatial thinking: The importance of frame of reference.* Cambridge University Press.

35. Williams, K.W., (2004). *A Summary of Unmanned Aircraft Accident/incident Data: Human Factors Implications (Technical Report DOT/FAA/AM-04/24).* U.S. Department of Transportation, Federal Aviation Administration, Office of Aerospace Me, Washington, DC.

36. Wilson, G.F. (2002a). *An Analysis of Mental Workload in Pilots during Flight Using Multiple Psychophysiological Measures.* The International Journal of Aviation Psychology, 12, 3-18.

37. Worringham, C.J., Beringer, D.B., (1998). *Directional stimulus-response compatibility: a test of three alternative principles.* Ergonomics 41 (6), 864e880.

38. Zupanc, C.M., Burgess-Limerick, R.J., Wallis, G., (2007). *Performance consequences of alternating directional control-response compatibility: evidence from a coal mine shuttle car simulator.* Human Factors 49 (4), 629-636.