Estimating Human Reliabilities for Maritime Control Systems

Mike Tainsh, PhD, FIEHF[1]

[1]BAE Systems Maritime, Frimley

ABSTRACT

Users of human reliability estimates may experience difficulties in use and acceptance, when addressing applications of maritime control systems. The research literature is reviewed with the conclusion that an additional set of reliability estimates is required. The method of paired comparisons was used to estimate the reliability of users' task performance for maritime applications when developing control systems. The reliabilities for the tasks are presented in rank order of likelihood of error. The information from this work is being used beneficially in support of performance assessments and safety justifications in development programmes.

KEYWORDS

Maritime control system, human reliability estimates, rank order of likelihood of error

Introduction

Human Reliability Assessment (HRA) has been an ergonomics analysis technique in use for some decades (HSE, 2009). It is commonly associated with hazardous situations such as those which may be found in the nuclear and related industries. Some techniques are now regarded in some quarters as de facto standards and applied widely, while failing to take full consideration of the constraints associated with them.

The constraints encountered during applications work in the defence industries, including maritime, may include:

(a) Ease of use – a "standard technique" may be identified for a project but when it is used, it is found difficult to map the set of tasks under consideration within the application onto the "standard set". This set of difficulties uses up resources and degrades validity.
(b) Acceptance – interpretation by Users. The results from a standard set of tasks can then be difficult to understand by users who have well organised understandings of their own roles and tasks within operational scenarios. This can lead to the results being rejected even if they are valid. The lack of acceptance can be very damaging to the ergonomic contribution to a project.
(c) Validity of application – If the mapping of the "standard" set of tasks, or contexts on to the system's task is not "one-to-one" then there may be a loss of validity which is unknowable in during system assessment. This is unacceptable when carrying out performance and safety investigations.

The purpose of this work is to develop a technique/information that aids “ease of use” and “acceptance”, and consequently enhances the validity of the assessments carried out in our maritime control system developments.

Current Research

Human Reliability (HR) is quantified using measures or categories of likelihood in an analogous approach to that used for non-human applications (Kirwan 1994). However, there are now next generation techniques being developed as exemplified by Yesim Kop Naskali, Tuncay Gurbuz and Y. Esra Albayrak (2019) and Huimin Ye and Wei Zheng (2016) but these techniques do not yet have results that are sufficiently mature for development work.

There is a substantial body of work on HR within complex control systems. The work includes Jianxin Huang, and Yaqin Bian (2011) and Yundong Guo, Youchao Sun , Xiufang Yang, and Zongpeng Wang (2019) in avionics and flight safety, and Abdelmoula Ait Allal, Khalifa Mansouri, Mohammed Qbadou, and Mohamed Youssfi (2017) on autonomous ships. These studies are concerned with the better exercise of control in their chosen applications and employ established estimates of performance. It is not clear how these might help “ease use” or “acceptance” in the maritime application now or in the future, apart from indicating progress in allied fields.

One approach within some applied studies appears to support the use of the current set of probabilistic estimates provided by HEART (Williams 1986) and seek the development of quantification of Error Producing Conditions (EPCs) as a way ahead. These studies include Xing Pan, Congjiao He, Tianjian Wen (2019) on task context and Subeer Rangra , Mohamed Sallak, Walter Schon, and Frederic Vanderhaegen (2017) on railway operations. Again, this work appears unable to help the current maritime issues, which do not include problems of understanding EPCs.

All the above, working on current complex control systems, appear to believe that the sets of tasks within the current HR assessment tools, such as HEART, are essentially “fit for purpose” but need some adjustments in the means of calculation techniques, or the associated EPCs to be applied successfully.

The outcome of this brief review suggests that current work on this technique does not cover those constraints which are currently encountered in maritime control systems. These approaches cannot help to solve the problems of “ease of use” and “acceptance” of the set of tasks or their associated reliabilities within the maritime control systems.

But most important, the use of the above work, leaves open the issue of validity which arises from the inability to map currently available generic task categories onto multi-layered control system tasks.

Hence the aim of this work is specifically to generate a set of tasks and human performance reliability estimates.

The research issues

The initial research issue is well established. There are two development options:

(a) Use an established bank of error rates which are known for a specific application.
(b) Create an application-specific set of error rates for use in a system development.

There is no established set of error rates for current maritime control systems. So, we need to use an established technique appropriate to this class of system that can generate the estimates required.

Aim of the Study

The overall aim of this work is to develop a HRA technique and data for use in the assessment of the system performance effectiveness or its capability to meet safety goals. A previous report (Tainsh, 2020) has shown, how the roles and tasks of teams of users working in computer-based control systems can be represented for assessment purposes. However, the appropriate HR information for maritime tasks was missing.

This paper focuses on the details of a HRA technique: the development of a tailored set of user error rates which can be employed within the assessment framework already constructed to assess the relative benefits of designs in support of performance and safety goals (Tainsh 2020). This work builds on current techniques rather than starting anew.

The Starting Point

Work in the field of HRA has traditionally started from one of two points:

(a) An understanding of the tasks to be carried out and the system context.
(b) The purpose to which the assessments will be used.

The maritime control systems used for naval operations are well known at a high level, and the context of their operations is widely reported.

However, the need to show that performance and safety cases will be fulfilled as a result of the design and implementation phases of a project means that HRA values are required for early work to help identify items of high risk and ensure their mitigation.

The Selection of Control System Tasks

The tasks listed here (Tainsh 2020) reflect the organisation, control structure and individual task content.

Organisational structure

The roles with equipment for maritime control systems are typically dived into three layers (Figure 1) (policy/senior management is combined here) and roles allocated accordingly. The roles to be addressed in this study are:

(a) Sensor team
(b) Collation of information/coordination team
(c) Decision-making
(d) Effectors team

Details of the platform and machinery control (the effectors) are not covered in detail here as these tasks and activities are procedural and more easily covered by techniques such as HEART (Williams 1986).

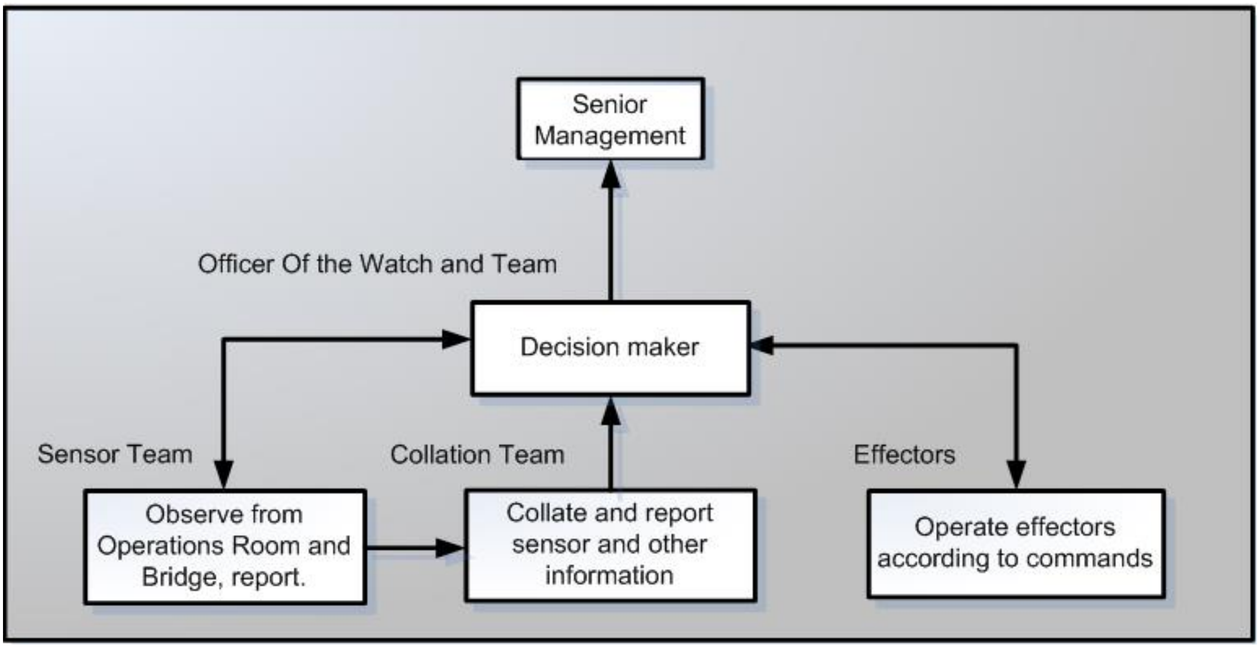

Figure 1: Three layered description of maritime control system

Control Structure

The control structure has the Officer Of the Watch as the decision maker controlling the movement of the vessel or changes in machinery state. Feedback on the consequences of control actions comes via the sensors detecting and tracking own and other vessels.

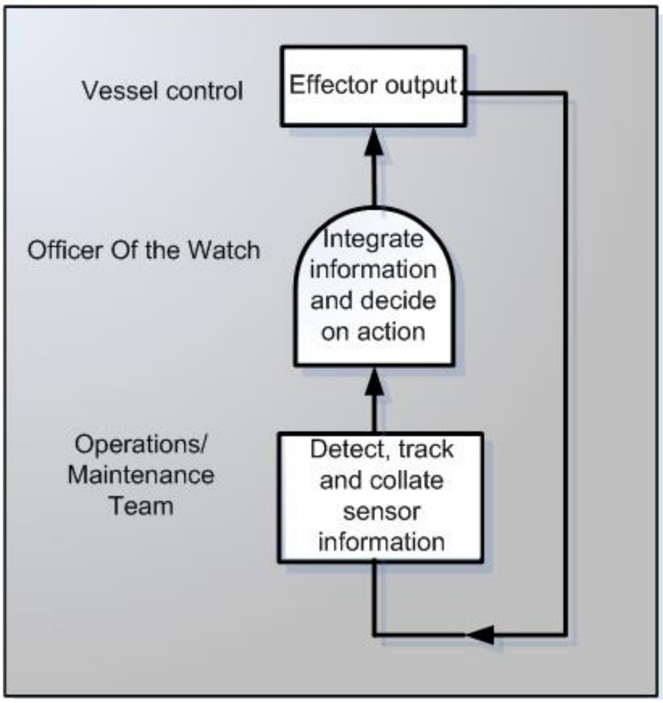

Figure 2: OOW Tasks represented as high-level control loop

Individual task context

All work in control system teams is carried out by users who understand its context. The selection of tasks for inclusion within the technique must reflect both the role and context of the work. The tasks must be general i.e. widely recognisable across platforms and systems.

The selection of the set of general tasks (Table 1) was conducted by ergonomics and engineering maritime specialists and Royal Navy personnel. It was also monitored throughout the comparison procedure to ensure that no general task was missing, nor were any redundant.

The tasks within Table 1 reflect those carried out within maritime control systems, taking account of the context of their actions.

	Context		
	Act on current information	Take account of history	Anticipate future events
Sensor systems	Initial detection (h), (i)	Tracking (e)	n/a
Information Organisation/coordination	Collation of information (g)	Classification/identification of contacts (a)	n/a
Decision-making	Interpreting information (b)	Checking against history (c), (d), (f)	Assessing potential outcomes (c), (d)
Effectors	Single action or operation (j)	Routine drills/sequences of actions (j)	n/a

Table 1 – General control tasks and context – the tasks map to Table 2 as indicated.

The mapping of the tasks from Table 1 to Table 2 is shown with the aid of identifiers (a) through to (i).

In addition, we consider the role of equipment fit and whether the person is using an item of equipment or not. For example, on the Bridge the sensor system may simply be the person's unaided vision

Hence, it is required to have a set of tasks that covers not the conditions shown in Table 1 but also the possibility of working with computer based equipment and also directly with an external environment e.g. looking out from the platform.

The estimation technique requires the inclusion of "calibration tasks" which enable the calculations to conclude (Kirwin, 1994). Two "calibration tasks" were identified. The operation of a valve was included as the first one of the set of tasks as it is a control systems task which can be easily identified. The second task was not identified until after the completion of the comparative judgements when it was identified as difficult/prone to error and could be assigned an agreed error rate based on current literature.

A pack of briefing material, including an uncompleted Comparative Judgement Table, was developed to enable the process of "Comparative Judgements" to be executed.

Method

The procedure followed is based on experience within the nuclear industry (Hunns and Daniels, 1980) and used the statistical process and information from Kirwin (1994). Their work uses Thurstone's "comparative judgements technique". While other techniques are available, it is known that military users find this technique understandable and acceptable (Keane and Avis, 2017)).

The selection of the Subject Matter Experts (SMEs) was carefully controlled to ensure that all had the background experience of the tasks during operation. The eight participants were from an operational or maintainer background, with an appropriate spread of ranks and specialist skills/experience, to ensure detailed knowledge of all the tasks including how well they are performed at sea. All had substantial lengths of service within their ranks and roles.

All the SMEs were volunteers who were available within easy communication.

In summary the comparison process was carried out in the following stages:

(a) The Paired Comparison Task was carried out with each of the experts using the "Comparison Table". This was always completed in a single session.
(b) The author asked the participants to make the comparisons. They all followed the same process. The comparison table was described to the participants using the tasks in Table 2. The comparisons were made starting at the top row and working across the columns. The instruction was:" Compare the task in each row with the task in each column so that the task more likely to have a lower error rate is assigned a one and the one less likely to have a low error rate is assigned a zero". The allocation of ones and zeros was recorded by the author.
(c) There was no communication between the participants.

The data analysis progressed through four stages

(a) The data from each participant was recorded on a data sheet and a "Raw Frequency Matrix" compiled.
(b) The "Proportion Matrix", was calculated, and the "Transformation X Matrix" derived
(c) The "Column-Difference Z Matrix" was derived.
(d) The scale values were calculated using the calibration points. The two calibration tasks, using HEART categories for unreliability are operation of valve (completely familiar routine task, 0.0004) and classification of contact (complex task requiring high level of comprehension and skill 0.16).

Results

The judgements of the participants using the paired comparison table in combination with the data analysis, enabled Table 2 to be constructed.

Table 2 shows the ranking of error rates associated with the set of tasks. This summarises the tasks and the set of information that is available using the comparative judgements technique. There is insufficient space here to provide the additional information that has been obtained.

After the comparison values were calculated all the participants were made aware of the results to check that the results were meaningful i.e. corresponded with their understanding of task performance during their execution of operations or maintainer roles. This contributed to our understanding of the “ease of use”. of the results and their likely “acceptance”

Outcome

The results of this programme of work are summarised in Table 2.

Identifier	User Task	Rank
(a)	Classification of contact	1
(b)	Interpreting and using tactical picture information	2
(c)	Officer Of the Watch, decision making within the vessel	3
(d)	Navigation, planning progress using tactical picture	4
(e)	Contact tracking with history	5
(f)	Officer Of the Watch decision making on the Bridge	6
(g)	Collation of contact information	7
(h)	Initial sensor detection of contact	8
(i)	Lookout/Bridge	9
(j)	Operation of valve	10

Table 2: Tasks ranked in order of likely error rate – most likely at the top

Conclusions

This work has depended on much that has gone before. The use of relatively a small-scale assessment based on the well-established ergonomics technique of “paired comparisons” as used in the nuclear and similar industries has enabled the identification of a set of general tasks that cover the requirement to describe the operations in Fig 1 and Fig 2, and the HR estimates that have been summarised here.

The use of a well-developed process has enabled the generation of understandable, easy to use and acceptable HR estimates. No quantified evidence on validity is yet available. However, reports from initial examples of use have been favourable.

These results will be useful for operational and safety modelling and risk mitigation where it is essential to understand whether task designs meet criteria such as Risks At Operationally Acceptable Levels (RAOAL) for operations i.e. levels at which users believe that they can cope and mitigate their consequences during operations to attain a goal, As Low As Reasonably Practicable (ALARP) for safety.

References

Abdelmoula Ait Allal, Khalifa Mansouri, Mohammed Qbadou, Mohamed Youssfi (2017) Task Human Reliability Analysis for a Safe Operation of Autonomous Ship. 2nd International Conference on System Reliability and Safety.

HSE/HSL (2009) Review of Human Reliability Assessment Methods.

Huimin Ye and Wei Zheng (2016) A Human Reliability Analysis Method based on Cognitive Process Model for Risk Assessment. IEEE International Conference on Intelligent Rail Transportation.

Hunns, D. and Daniels B.K. (1980) Paired comparisons and estimates of failure likelihood. Design Studies. Vol 2, No 1.

Jianxin Huang, and Yaqin Bian (2011) Human Reliability Analysis in Complicated Armament Equipment System. IEEE Conference Quality, Reliability, Risk, Maintenance and Safety Engineering.

Keane D. and Avis C. (2017) A Tool to Generate 'HF Meaningfulness' in the Design and Development of Armoured Fighting Vehicles. Proceedings of CIEHF Annual Conference.

Kirwin B (1994) A Guide to Practical Human Reliability Assessment. Taylor and Francis. London

Subeer Rangra, Mohamed Sallak, Walter Schon, and Frederic Vanderhaegen (2017) A Graphical Model Based on Performance Shaping Factors for Assessing Human Reliability IEEE Transaction on Reliability, Vol 66(4).

Tainsh (2020) Comparison of safety system designs using Risk Assessment. Proceedings of CIEHF Annual Conference.

Thurstone, L.L. (1927). The method of paired comparisons for social values. Journal of Abnormal and Social Psychology, 21, 384-400.

Williams, J C (1986) A proposed method for assessing and reducing human error. In 9th Advances in Reliability Technology Symposium, University of Bradford.

Xing Pan, Congjiao He, Tianjian Wen (2014) A Review of Factor Modification Methods in Human Reliability Analysis. IEEE International Conference on Reliability Maintainability and Safety.

Yundong Guo, Youchao Sun, Xiufang Yang, and Zongpeng Wang (2019) Flight Safety Assessment Based on a Modified Human Reliability Quantification Method. International Journal of Aerospace Engineering.

Yesim Kop Naskali, Tuncay Gurbuz and Y. Esra Albayrak (2019) Human Reliability Assessment in Healthcare Operations Using Fuzzy Cognitive Maps. Journal of Mult.-Valued Logic and Soft Computing. Vol 32 pp 578-86.