NARIDAS – EVALUATION OF A RISK ASSESSMENT SYSTEM FOR THE SHIP'S BRIDGE

B Gauss and **M Rötting**, Berlin University of Technology, Department of Human-Machine Systems, Germany
D Kersandt, AVECS Corporation AG, Germany

SUMMARY

The Navigational Risk Detection and Assessment System (NARIDAS) is a novel approach to a task-oriented integration and assessment of nautical data on the ship's bridge. Based on about 100 physical and technical input parameters of the navigation process, NARIDAS performs an online calculation of the current situation's navigational risk on eight dimensions. With a bar chart of the eight risk values, NARIDAS provides a comprehensive overview of the current risks to support situation awareness of the bridge team. The focus of this paper is on the evaluation of NARIDAS during the development process. Evaluation is conducted with practitioners and nautical experts, and addresses three levels: (1) risk model validity, (2) design of graphical user interface, (3) system effectiveness. Two evaluation studies were conducted. Study I was carried out with a static functional prototype, addressing levels (1) and (2). For study II, a fully-functioning prototype was implemented in a ship-handling simulator to investigate level (3). Positive results were obtained on all three levels, indicating that NARIDAS provides a valid model for the situational risks of ship navigation, and a promising tool for enhancing situational risk awareness of the bridge team.

1. INTRODUCTION

Accident analyses show that human error is a dominant factor in about 80% of maritime accidents. Many of these human error accidents are attributed to failures of situation awareness [1]. It has been argued that it is often not adequate to attribute the main causes of accidents in complex human-machine systems like ships to active failures of the operators 'at the sharp end' (here: on the ship's bridge). Instead, it appears more promising for the prevention of future accidents to investigate the latent failures in the system 'at the blunt end' (design, organisation, management etc.) [2]. In hindsight, for the accident researcher, it might be rather easy to detect operator errors in the chain of events resulting in an accident. But in many cases, things were less clear for the bridge team during the event: 'Errors do not look like errors at the time they are perpetrated, and the accidents that are caused by them look impossible beforehand' [3].

However, there is no doubt that situation awareness of the bridge team was insufficient in many severe marine accidents. Situation awareness can be defined as 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' [4]. If we have identified situation awareness failures as an important cause of accident, we should ask *why* bridge crews lose situation awareness. The answer to this question will enable us to find out what can be done that situation awareness is enhanced or at least maintained in critical situations.

2. THE ROOT CAUSE: DATA OVERLOAD

The overabundance of data provided by the technical systems on the ship's bridge is a crucial factor that makes the task of maintaining situation awareness difficult. Watch officers are confronted with data overload, caused by the ever-growing number of displays, screens and navigation devices on the bridge.

Data overload is considered a significant problem in many domains of human-machine systems: 'The ubiquitous computerisation of the modern world has enormously advanced our ability to collect, transmit and transform data, producing unprecedented levels of access to data. However, our ability to interpret this avalanche of data, i.e., to extract meaning from artificial fields of data, has expanded much more slowly, if at all. In studies across multiple settings, we find that practitioners are bombarded with computer-processed data, especially when anomalies occur. We find users lost in massive networks of computer-based displays, options and modes' [5].

On the ship's bridge, the poor design of user-interfaces aggravates this problem. The integration of different navigation aids is insufficient, and there are no common standards for user-interfaces on the ship's bridge [6]. Screens tend to be complex, packed with numeric information, and difficult to operate correctly. In addition, the number of alarms is often confusing.

Generally, automated systems can support four different stages of human information processing and behaviour [7] (Figure 1). Today shipboard automation concerns mainly the first and the last stage: information acquisition with all kinds of sensors and the related displays, devices or systems (e.g., ARPA, GPS, AIS) and action implementation (e.g., autopilot, track control).

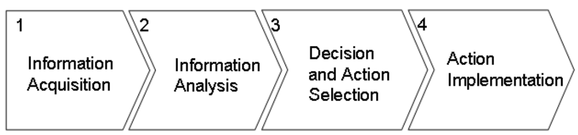

Figure 1: Stages of human information processing

In consequence, the prevailing approach to modern aids for ship navigation has reached its limit. More and more information is acquired and presented by technological systems, but the tasks to assess the information from multiple sources and to decide what to do next remain with the human operators. The bridge team cannot profit any more from the very fast and accurate numerical description of the ship navigation process provided by modern computer-based systems, because there is too much data available.

3. THE SOLUTION: NARIDAS

A possible solution to the problem of data overload, caused by the continuous expansion of 'information acquisition systems', is the development of support systems for the cognitive processing stage of information analysis. On this stage, the information acquired on the first stage is integrated by relating it to the current goals. The operators extract the *meaning* of the information in their task environment for decision and action selection.

An important semantic category at this stage is the concept of *risk*. Risk can be defined as the anticipation of an event with negative consequences. In dynamic human-machine systems, subjective risk assessments are directly linked to decision making and action. If subjective risk is too high, the operator will change his or her plan and take adjusting actions to reduce risk to an acceptable level. Of course, it is crucial for adequate decision making that risk is assessed correctly, i.e., that the operators' subjective risk reflects the situation's actual or 'objective' risk. Thus, a risk assessment system could support the cognitive processing stage of information analysis in order to overcome the data overload problem. In addition to the raw sensor data of information acquisition systems, a risk assessment system offers a task-oriented integration of the acquired information.

For ship navigation, the *Navigational Risk Detection and Assessment System* (NARIDAS) is a novel approach to support integration of nautical data by dynamic risk assessments. The basis of NARIDAS is the breakdown of the navigation process into eight task dimensions [11]:

- COLLISION AVOIDANCE (COL): pass other ships or objects safely;
- ANTI-GROUNDING (GRD): adjust own ship's speed to the natural conditions;
- TRACK KEEPING (TRA): keep track and consider manoeuvring area;
- TRAFFIC (TRF): account for characteristics and density of traffic;
- BRIDGE MANNING (MAN): consider the condition of the bridge crew;
- ENVIRONMENT (ENV): account for the meteorological and hydrological conditions;
- ENGINE/WHEEL (ENG): consider the state of propulsion and rudder engines;
- ECONOMY (ECO): comply with the economic criteria of the voyage.

For each of these task dimensions, NARIDAS calculates the corresponding risk by means of knowledge-based and rule-based procedures. In a first step, about 100 technical or physical input parameters – that are continuously updated from various sources (e.g., ARPA, electronic chart, integrated navigation system) – are processed by crisp mathematical algorithms for nautical calculations. In doing so, the input parameters are integrated into 24 higher-order variables. These higher-order variables are further processed with fuzzy algorithms comparing their current values with standard values for 'good seamanship' to obtain the eight navigational risk values on a scale from 0='no risk' to 1='accident'.

As an example, the COLLISION risk is assessed by combining closest point of approach (CPA), time to closest point of approach (TCPA), and other characteristics of all radar targets, based on ARPA data. In addition to the radar image, ARPA displays present a multitude of numeric information for each target (e.g., bearing, heading, course, speed, CPA, TCPA). NARIDAS detects the three most dangerous ARPA targets and calculates one corresponding risk value for the current situation. So, ARPA information is reduced drastically and integrated into one meaningful category, 'collision risk', which is directly related to the tasks of the human operators.

On the graphical user interface of NARIDAS, the eight situational risk values are displayed in a bar chart (Figure 2). This comprehensive display allows for an assessment of the situational risks of the navigation process at a glance. Also, the system offers access to more detailed explanations, so the users can check the reasons behind the system's risk assessments.

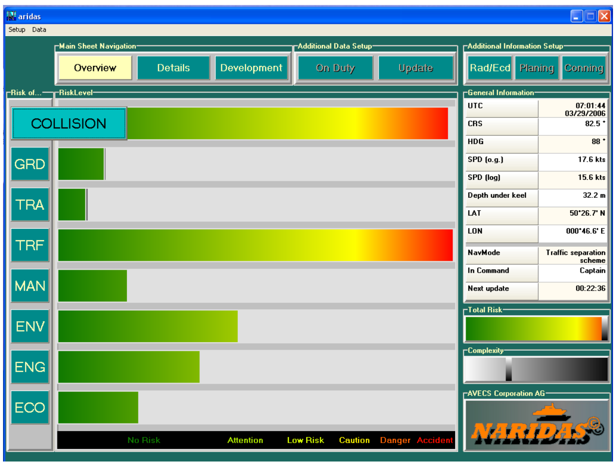

Figure 2: NARIDAS Graphical User Interface (GUI)

Since navigational risks are context-specific, the NARIDAS knowledge-base is customised on three different levels: (1) *long term*: to the particular ship (manoeuvring properties, engine characteristics etc.), (2) *medium term*: to the voyage plan (way points, estimated time of arrival etc.), and (3) *short term*: to the current sea area. For the latter, NARIDAS distinguishes between six different 'navigation modes':

- Coastal waters
- Approach
- Traffic separation scheme
- Fairway
- Open sea
- At anchor

For each navigation mode, a specific set of standard 'good seamanship' values is activated. As an example, NARIDAS permits a smaller CPA in restricted waters than in open sea for COLLISION (e.g., a CPA of 1 nm is rated 'high risk' in open sea, while it is rated 'OK' in a traffic separation scheme).

4. DEVELOPMENT PROCESS AND EVALUATION FRAMEWORK

In many cases, the development of maritime automation and support systems is technology-driven. The application of user-centred processes for system development is not common in the maritime domain. As a result, many high-tech navigation aids are notorious for their poor usability [8]. Practitioners find the situation on modern ships bridges an 'ergonomic nightmare' [9].

Recently, improvements for the design process have been proposed by applying human factors engineering approaches, e.g., for a military integrated bridge system [10]. The main characteristics of the proposed human factors methodology consist in (1) the use of mock-ups and prototypes, and (2) the involvement of the future operators in evaluation studies. For the NARIDAS development process, a '*parallel-iterative approach'* of system development is applied:

- *Parallel*: From early stages of the development process, we work on technological and human aspects of the system in parallel.
- *Iterative*: The match of these aspects is controlled in iterative evaluation loops with prototypes and the participation of practitioners.

This approach requires the interdisciplinary cooperation of system developers, domain experts and human factors specialists. For the rest of this paper, we will focus on the evaluation process as the key task from the human factors point of view. The main objective of the evaluation is to gather information for the improvement of the system. Evaluation in the NARIDAS development process can be assigned to an 'evaluation pyramid' of three levels (Figure 3). On the basic level, the validity of the NARIDAS risk model, is verified. Secondly, the design of the graphical user interface (GUI) is reviewed. Finally, the effectiveness of the complete system is evaluated.

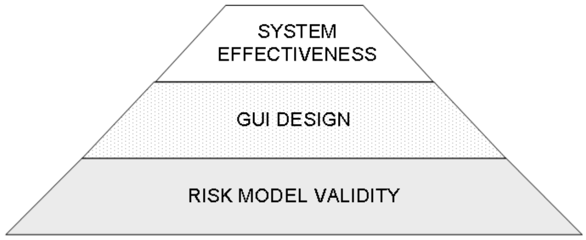

Figure 3: Evaluation pyramid

4.1 RISK MODEL VALIDITY

NARIDAS was invented by one of the authors, Diethard Kersandt, on the basis of his vast experiences in practice as Master and Nautical Officer as well as in the academy as nautical instructor and accident researcher. He designed and adjusted the NARIDAS knowledge base, i.e., the algorithms for risk calculation, in several years of development work. The evaluation objective at the first level is to check how well the algorithms represent the risk assessments of other nautical experts. Of course, it is an essential prerequisite for acceptance and effectiveness of NARIDAS that the risk algorithms reflect the common view on risk and not just the personal opinion of a single expert.

4.2 GUI DESIGN

The most important question at this level is how the risk values should be displayed to provide an optimal overview of the situation. Also the presentation of the additional information (e.g., the details of the risk calculations, explanation components), the menu structure, and general usability criteria (e.g., error tolerance, learnability, acceptance) of the GUI have to be evaluated.

4.3 SYSTEM EFFECTIVENESS

For sure, the demonstration of the system's effectiveness is crucial for success. Why should the ship-owner buy a system, why should the nautical officer use a system, as long as its effectiveness is not clear? Unfortunately, the effectiveness of a risk assessment system is difficult to prove. On the one hand, as we have discussed at the beginning of this paper, it seems obvious that support on the cognitive stage of information analysis should contribute to reduce data overload on the bridge. On the other hand, the benefits of a cognitive support system might appear to be rather subtle, compared to 'classical' navigation aids like radar. Radar makes objects visible that would be invisible without it in conditions of fog or darkness. A risk assessment system makes risk visible. Before, risk has never been visible, and we have been navigating safely for thousands of years. We have done so without radar, too. What exactly will we gain if we can see the navigational risks?

5. EVALUATION PROCEDURE AND RESULTS

Until now, two studies have been conducted in the NARIDAS evaluation process. Study I addressed the two basic levels of the evaluation pyramid. Study II investigated the top level.

5.1 STUDY I: RISK MODEL AND GUI DESIGN

Study I was carried out in two rounds with a functional NARIDAS prototype, which presented the GUI, and contained the nautical data and risk values for several pre-defined static traffic scenes. The objectives of the study were (1) to compare the NARIDAS risk values with risk judgements of nautical experts and (2) to enquire the experts' opinions about the GUI and the overall concept of this nautical risk assessment system.

5.1 (a) Procedure

Participants were 16 nautical experts (masters, mates, final-year students) of German nationality. All of them were between 25 and 60 year-old men with nautical experience on board of large vessels world-wide. The study was conducted in individual trials. After an introduction to NARIDAS, 14 static traffic scenes were presented to the expert. These scenes represented a broad range of different navigational requirements (e.g. passing Straight of Gibraltar; approaching port of Livorno; open sea) and environmental conditions.

For each scene, the experts received data about own ship characteristics (pilot card), traffic situation and sea area (screenshots of radar and electronic chart), and environmental data (wind, waves, visibility etc). Experts were instructed to judge the navigational risks of the traffic scene on the eight dimensions. After the risk assessment a computer screen with the functional prototype was switched on, so that the experts could explore the system and compare their own risk assessments with the NARIDAS values. During risk assessment and system exploration, experts were asked to think aloud. Verbal data was recorded, transcribed and analysed qualitatively. After completion of the risk judgements, a detailed usability questionnaire with rating questions was administered. At the end of the trial a short structured interview was held on the experts' opinions about NARIDAS.

5.1 (b) Results

Over all traffic scenes and risk dimensions, experts' judgements and NARIDAS values were highly consistent (Cronbach's Alpha between .89 and .94). Figure 4 shows examples of experts' (means) and NARIDAS risk assessments in 3 different scenes. In a sensitivity analysis, rates of 'misses' and 'false alarms' were determined. A miss was defined as a case if >50% of the experts assessed a risk as 'dangerous' (>.80) and NARIDAS assessed the risk as 'not dangerous' (<.60). A false alarm was defined as a case if NARIDAS assessed a risk as 'dangerous' and >50% of the experts assessed the risk as 'not dangerous'. With 1.5% (1st round) and 7.1% (2nd round) of false alarms, and 0.9% of misses (both rounds) for a total number of 112 cases (= 14 scenes * 8 risk values), sensitivity of NARIDAS was high.

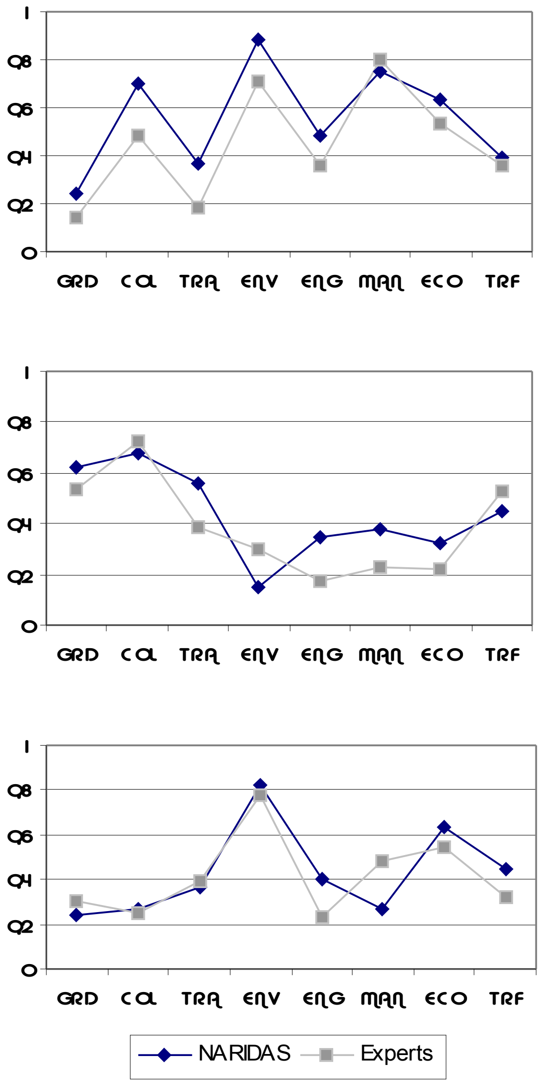

Figure 4: Experts' (means) and NARIDAS risk assessments in three different traffic scenes

In the questionnaire, the GUI was rated very positive. Participants judged the NARIDAS interface as clearly designed and easy to use. Overall usability of the system, assessed on a 10-items-scale (e.g., 'NARIDAS is a reliable system', 'NARIDAS would enhance the safety of navigation'), achieved 73.2 of 100 points. Also the qualitative data (think-aloud protocols, interviews) showed that the experts considered NARIDAS as a useful support to ship navigation.

5.2 STUDY II: SYSTEM EFFECTIVENESS

For study II, a fully-functioning NARIDAS prototype was implemented in the ship-handling simulator in Elsfleth, Lower Saxony (Figure 5). The objectives of the study were to investigate the effects of NARIDAS on situational risk awareness and navigation performance of the bridge team during a simulated voyage.

Figure 5: NARIDAS in the Elsfleth simulator

5.2 (a) Procedure

NARIDAS was connected to the simulator network, so it was calculating the risks online during the whole voyage. This was the first test of the system under dynamic conditions, after a basic prototype had been installed on board a cruise ship in 1996.

Participants were 23 nautical students in the final year of their studies (all men; age between 21 and 48 years, mean=28 years). They were grouped into 11 bridge teams, each team consisting of one 'Master' and one or two 'Watch Officers'. A traffic scenario in the English Channel of 80 minutes was constructed with high traffic density and rather unpleasant environmental conditions (4m swell from 220°, 30kn wind from 180°, 2.5kn current from 50°). Own ship was a container vessel travelling from Cadiz to Rotterdam.

A simple one-factor experimental design was realised with 'NARIDAS support' as independent variable, which was varied within teams. Each team travelled one 40-minutes section of the trip with NARIDAS, the other 40-minutes section without NARIDAS (i.e., NARIDAS display was switched off). The sequence of sections with and without NARIDAS was balanced between teams. Dependent variables were assessed with a combination of different methods. After each section, rating questionnaires were applied to assess situational risk awareness (SRA) and navigation performance (self-ratings by the subjects and assessment of the teams by an experienced instructor). Furthermore, SRA was measured with an online-test, 3 times during each voyage section (after 15, 25 and 35 minutes). For this test, the 'Master' received a phone call from the experimenter. He was asked to report the three most dangerous risks at the particular moment, and to rate these risks on a scale from 0 to 100. The answers were recorded, and categorised ex post to the NARIDAS risk dimensions for analysis. As an additional indicator for navigation performance, NARIDAS risk values were recorded during the whole trip.

5.2 (b) Results

In the SRA online-test, subjects had higher risk awareness in the sections travelled with NARIDAS support (Figure 6). In particular, more collision risks (i.e., dangerous radar targets) were reported by the participants. The difference between the sections with and without NARIDAS is statistically significant (Wilcoxon-Test, $p<.01$). Results also show that only three of the eight risk dimensions (collision, environment, and traffic) were rated 'dangerous' during the test. This indicates that overall complexity of the traffic scenario was rather low. The teams were able to handle the requirements of this simulator exercise without major problems.

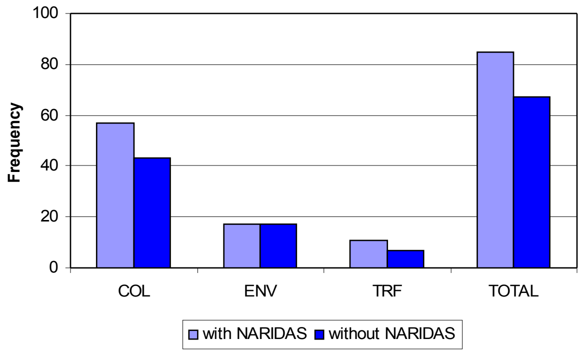

Figure 6: Reported risks in the SRA online-test

Analysis of navigation performance showed that with NARIDAS, a higher risk of collision (the most important risk dimension in the scenario) was associated with better SRA and navigation performance ratings by the instructor. Without NARIDAS, a higher risk of collision was associated with a poorer instructor rating (Table 1).

Table 1: Correlations between instructor ratings and recorded collision risk (Spearman-Rho; *p<.05)

		Collision Risk	
		With NARIDAS	Without NARIDAS
Instructor Ratings	Navigation Performance	.48	-.59*
	SRA	.68*	-.45

This result suggests that NARIDAS can contribute to a better handling of high risks. If a high risk is taken consciously (i.e., with a high SRA, supported by a risk assessment system), navigation performance is good, and the situation remains under control. In contrast, if the navigators take a high risk without recognising it (lower SRA, no support), navigation performance becomes unstable.

Again, usability of NARIDAS was rated positive by the participants, and their acceptance of the system was high. In an overall judgement, 19 participants rated NARIDAS as 'good' or 'very good', the other 4 participants as 'neither good nor bad'. There were no negative judgements on this novel system.

6. DISCUSSION

In the two empirical studies, results were encouraging on all three levels of evaluation. In study I, the risk values calculated by NARIDAS matched very well with the risk judgements of nautical experts. This result indicates high validity of the NARIDAS risk model. Furthermore, these findings imply that there is a common view on the navigational risks among nautical experts, and this common view can be modelled by a combination of mathematical and fuzzy-set algorithms. However, it should also be noted that consistency of risk assessments between the experts and NARIDAS, as well as inter-individual consistency between the different experts, is high but not perfect. If we use more abstract concepts like risk, we will be confronted with a higher degree of uncertainty than with crisp technical or physical parameters. In complex, dynamic processes like ship navigation, human operators will always have to cope with uncertainty. The concept of risk makes uncertainty measurable and visible. The positive expert ratings on usability and acceptance suggest that practitioners believe they will profit from the display of risks by NARIDAS, despite the residual fuzziness of the risk concept.

For study II, NARIDAS was successfully implemented in the full-mission ship-handling simulator Elsfleth, so the system's operational capability could be demonstrated online in a dynamic setting. Experimental comparison showed positive effects of NARIDAS on situational risk awareness and navigation performance, even though the voyage scenario realised for the simulator study resulted to be not extraordinarily challenging for the well-trained participants. In the future, NARIDAS should be tested under more tricky conditions, e.g. a slowly evolving emergency scenario in a simulator exercise of several hours. We assume that the benefits of NARIDAS should appear even clearer if the bridge team had to switch unexpectedly from operational routine to a peak workload situation. In study II, workload was rather moderate without major variations during the exercise, reflecting an everyday's working scenario.

Furthermore, the measurement of navigational risks by NARIDAS offers perspectives for various applications beyond the use as support tool for the bridge. In the ship-handling simulator, NARIDAS could provide online training feedback for the students as well as standardised assessments of navigation performance. Last but not least, dynamic risk assessments could be integrated into voyage data replay systems. So, incident and accident analyses would profit from risk profiles of critical situations, e.g. to determine the 'point of no return'.

7. CONCLUSIONS

In conclusion, the results of the evaluation studies show that NARIDAS provides a valid model for the risks of ship navigation, and a promising tool for reducing data overload and enhancing situational risk awareness of the bridge team. In both studies, usability of NARIDAS and acceptance by the practitioners were high.

Besides the presentation of a novel support system, the purpose of this paper is to give an example for the role of evaluation and for useful procedures within a parallel-iterative approach to system development in the maritime domain. The objective of this methodology is to achieve a user-centred design of technological systems, by involving the future users and considering human factors aspects from the very beginning of the design process.

8. ACKNOWLEDGEMENTS

The simulator study reported in this paper was kindly supported by Mr Klaus Damm and Mr Pawel Bednarz from the Elsfleth ship-handling simulator, Department of Maritime Studies at the University of Applied Sciences Oldenburg/ Ostfriesland/ Wilhelmshaven.

9. REFERENCES

References should be numbered in the text where they occur, with the number in square brackets [1] and a complete list given at the end of the paper:

1. Baker, C.C. & McCafferty, D.B., 'Accident Database Review of Human-Element Concerns: What Do the Results Mean for Classification?', *RINA Human Factors in Ship Design, Safety & Operation (pp. 65-71)*, 2005.
2. Reason, J., *Human Error*, Cambridge University Press, New York, 1990.
3. Wagenaar, W.A. & Groeneweg, J., 'Accidents at sea: Multiple causes and impossible consequences', *International Journal of Man-Machine Studies, 27, 587-598*, 1987.
4. Endsley, M.R., 'Toward a Theory of Situation Awareness in Dynamic Systems', *Human Factors, 37 (1), 32-64*, 1995.
5. Woods, D.D., Patterson, E.S. & Roth, M., 'Can We Ever Escape from Data Overload? A Cognitive

Systems Diagnosis', *Cognition, Technology and Work, 4, 22-36*, 2002.

6. Moreton, M.-B., *Human Factors on the Ship's Bridge*. PhD Thesis, Liverpool John Moore University, 2000.
7. Parasuraman, R., Sheridan, T.B. & Wickens, C.D., 'A Model for Types and Levels of Human Interaction with Automation', *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol. 30* (3), 286-297, 2000.
8. Lützhöft, M., *'The technology is great when it works. Maritime Technology and Human Integration on the Ship's Bridge'*. PhD Thesis, University of Linköping, 2004.
9. The Nautical Institute, *Alert! The International Maritime Human Element Bulletin,* Issue No. 3, April 2004, http://www.he-alert.org/ [visited 25 January, 2007]
10. Chouvy, N. & Pellen-Blin, M., 'Integration of Human Factors into the Design Process / Practical Applications of Human Factors Engineering', *RINA Human Factors in Ship Design, Safety & Operation (pp. 7-19)*, 2005.
11. Kersandt, D., 'ANRIS – Automated Navigation Risk Identification System – A knowledge based "Early Warning System" for avoiding the point of no return. *ISHFOB '95 The Influence of the Man-Machine-Interface on Safety of Navigation.* Proceedings of the International Symposium of Human Factors on Board. TÜV Rheinland.- Bremen, November 1995

10. AUTHORS' BIOGRAPHIES

Boris Gauss is a Research Assistant at Berlin University of Technology, Department of Human-Machine Systems, and holds a 'Diplom' (German MSc equivalent) in Psychology. He has been doing research on Human-Machine Systems for several years in various domains (e.g., air traffic control, aircraft evacuation, ship navigation). Currently he is preparing his PhD thesis on 'Situational Risk Awareness in Human-Machine Systems'.

Dr.-Ing. habil. Diethard Kersandt is consultant for maritime companies, working since more than 30 years on problems of human error and the development of support systems in this domain. He holds a 'Diplom' (German MSc equivalent) in Sea Transport and a Master's certificate for deep-sea navigation and went to sea for 8 years as Chief Mate and Nautical Officer on merchant ships. For more than 10 years he held the position of an Assistant Professor and Lecturer at the Maritime Academy in Warnemünde, Germany.

Prof. Dr.-Ing. Matthias Rötting holds the position of a University Professor and is the head of the Department of Human-Machine Systems, Berlin University of Technology. …