Interaction Design for Situation Awareness
－Eyetracking and Heuristics for Control Centers－

Henning BREUER[†][‡] Frank HEIDMANN[†] Monika WNUK± Ulrich RAAPE± Mitsuji MATSUMOTO[‡]

[†] University of Applied Sciences Potsdam, Faculty of Design, Pappelallee 8-9, 14469 Potsdam, Germany

[‡] Graduate School of Global Information and Telecommunication Studies, Waseda University, Waseda Bldg. 29-7, 1-3-10 Nishiwaseda, Shinjuku-ku, Tokyo, 169-0051 Japan 1-2-3 Yamada, Minato-ku, Tokyo, 105-0123 Japan

± German Remote Sensing Data Center, German Aerospace Center, Muenchner Str. 20, 82234 Wessling, Germany

E-mail: [†] {breuer;heidmann }@fh-potsdam.de ±{monika.wnuk;ulrich.raape}@dlr.de [‡] mmatsumoto@waseda.jp

Abstract Safety-critical applications require high degrees of usability and error tolerance. Supporting the situation awareness of operators has become a central topic in the design of control center applications. In order to evaluate the prototype of the control center screens of a decision support system for tsunami early warning we compiled a heuristic for expert evaluation and conducted eyetracking studies to explore the management of attention. We present the methods we applied, some results of the study, and exemplify improvements of the system under development.

Keyword: Safety-critical applications, tsunami early warning system, decision-support system, control centre, designing for situation awareness, heuristic, eyetracking, usability, error tolerance.

状況認識のためのインタラクションデザイン
コントロールセンターのための－アイトラッキングと発見的教授法

Henning BREUER[†][‡] Frank HEIDMANN[†] Monika WNUK± Ulrich RAAPE± Mitsuji MATSUMOTO

[†] University of Applied Sciences Potsdam, Faculty of Design, Pappelallee 8-9, 14469 Potsdam, Germany

[‡] Graduate School of Global Information and Telecommunication Studies, Waseda University, Waseda Bldg. 29-7, 1-3-10 Nishiwaseda, Shinjuku-ku, Tokyo, 169-0051 Japan 1-2-3 Yamada, Minato-ku, Tokyo, 105-0123 Japan

± German Remote Sensing Data Center, German Aerospace Center, Muenchner Str. 20, 82234 Wessling, Germany

E-mail: [†] {breuer;heidmann}@fh-potsdam.de ±{monika.wnuk;ulrich.raape}@dlr.de [‡] mmatsumoto@waseda.jp

アブストラクト: 安全性の局面を左右するアプリケーションは、ユーザビリティとエラー寛容性の高い程度を必要とします。オペレーターの状況認識を支えることは、コントロールセンターアプリケーションデザインの中心テーマになりました。津波早期警戒のために意志決定支援システムのコントロールセンタースクリーンのプロトタイプを評価するために、我々は専門家の評価のために発見的手法をコンパイルして、注意の管理を調査する研究のためにアイトラッキングを実行しました。我々が適用した方法を紹介します。いくつかは研究ニ起因しています。そして、開発中にシステムの改善を例証してください。

キーワード： 安全性重要なアプリケーション、津波早期警戒システム、意志決定支援システム、管制センター、状況認識のための設計、発見的手法、アイトラッキング、ユーザビリティ、エラーの寛容性。

1. Designing for Safety

Safety-critical applications include natural disaster early warning systems, or control centers for air and ship-traffic, or factory plants. Safety of such systems depends on properties of the system, and the users interacting with it. Even though safety-critical devices are often operated by trained personnel still natural exceptions to professional operation e.g. due to boredom, fatigue, breaks or contextual irritations need to be considered. Operational error e.g. due to the misinterpretation of signals and information, or the triggering of unintended system responses may have hazardous impacts. Facing this danger following standard usability guidelines and measures is insufficient.

Additional guidelines and techniques to evaluate interaction design have been proposed.

Situation awareness” has been identified as one essential psychological concept within the safe operation of complex systems. It deals with „Knowing what is going on so you can figure out what to do” [1]. „Situation awareness is the perception of 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“[2]. Endsley and colleagues differ between three levels of situation awareness (SA).

On the first and most basic level perception involves monitoring, cue detection and simple recognition. Level 1 SA then is an awareness of multiple situational elements (objects, events, people, systems, environmental factors) and their current states locations, conditions, modes, actions). On the second level comprehension involves pattern recognition, interpretation and evaluation; it produces Level 2 SA, an understanding of the overall meaning of the perceived elements - how they fit together as a whole, what kind of situation it is, what it means in terms of one's mission goals. On the third level projection involves anticipation and mental simulation; it produces Level 3 SA, an awareness of the likely evolution of the situation, its possible/probable future states and events.

In section 2 we present some dialogue principles to support users and guidelines to support and techniques to evaluate situation awareness in control centre.

Throughout the development and evaluation of the system we applied principles of user-centred design:

- Focus on “real” users and engage them early and continuously throughout the product life cycle.
- Validate UI requirements and designs by observing, measuring and recording end users.
- Design, prototype and develop UIs iteratively.
- Understand and design for “holistic” user XP

2. Project Background

The work presented here is embedded in the German-Indonesian Tsunami Early Warning System (GITEWS) project. GITEWS is funded by the German Federal Ministry of Education and Research (BMBF) to develop a Tsunami Early Warning System for the Indian Ocean in close cooperation with Indonesia, the country most prone for tsunamis in the whole Indian Ocean. The system integrates terrestrial observation networks of seismology and geodesy with marine measuring sensors, satellite technologies and pre-calculated simulation scenarios.

GITEWS is the German contribution to the Indonesian Tsunami Early Warning System InaTEWS. The GITEWS project uses sensor technologies to detect indicators or evidence for a tsunami, combining that information with up-to-date modelling techniques and integrating them in a newly developed Decision Support System. Combining a-priori knowledge, simulation runs and analysis results with real-time information from different types of sensors, the GITEWS Decision Support System (DSS) serve as a back-bone to allow an assessment for the tsunami threat at the earliest time possible and support the decision maker whether to issue a tsunami warning or not.

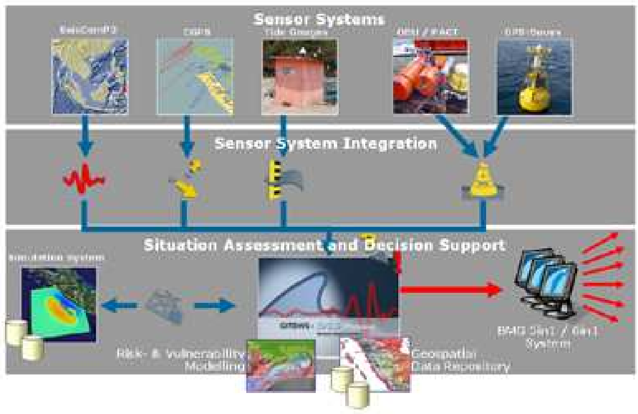

Image 1: Early Warning & Mitigation System Concept

Unlike classical decision support problems, the process of combining sensor and additional information, generating situation awareness and assessing and proposing decision options is a slowly evolving process. Due to the fact, that sensor information becomes available in a non-deterministic irregular sequence, initially with considerable uncertainties, in arbitrary order and with major information gaps, uncertainties will still be present when deadlines for warning decisions are reached.

Image 2: Test at the reference work station

Within a small research and consulting project we evaluated the Graphical User Interface of the GITEWS Tsunami Early Warning Decision Support System (DSS) Prototype using interface design heuristics and eyetracking analysis.

3. Situation Awareness and Dialogue Principles

Starting points in looking for suitable design principles are the internationally standardized 7 Dialogue Principles for Interactive Systems (ISO 9241-110) [4]:

Suitability for the task: An interactive system must support its user to achieve his tasks completely, correctly and with adequate effort.

Self-descriptiveness: A dialogue must enable the user to know at all times in which dialogue and which step he is in, which actions may be undertaken and how they may be undertaken.

Conformity with user expectations: An interactive system should be designed consistently and in line with user expectations and characteristics (like professional knowledge, experience and general conventions).

Error tolerance: An interactive system should keep user from making mistakes (e.g. applying security checks), but also support users corrective actions if a mistake occurred.

Controllability: An interactive system must be controllable by its user e.g. by offering an undo-function.

Suitability for individualization: An interactive system should be adaptable to the users characteristics and preferences and his tasks.

Suitability for learning: An interactive System should enable the user to learn interacting with the system.

Since we are dealing with a decision support system for experts focus was put on the first five of these principles. While they apply to interactive systems in general more specific guidelines have been proposed to support situation awareness of operators. The following principles from Endsley, Bolté & Jones [3] work on „Designing for Situation Awareness“ we added to our heuristic:

1. Organize information around goals
2. Present level 2 information directly in order to support an immediate comprehension of the situation
3. Provide assistance for level 3 situation awareness projections enabling anticipation of upcoming situations
4. Support global situation awareness presenting the „big picture“
5. Support trade-offs between goal-driven and data-driven processing
6. Make critical cues for schema activation salient
7. Take advantage of parallel processing capabilities
8. Use information filtering carefully

Additional principles provide guidelines for the design of representation of in the context of decision-support:

9. Explicitly identify missing information
10. Support sensor reliability assessment
11. Use data salience in support of uncertainty
12. Represent information timeliness
13. Support assessment of confidence in composite data
14. Support uncertainty management activities

From these principles we may derive some additional inspirations and recommendations for design of the early warning decision support system: Starting with operator goals: The elementary task of an operator (or even dispatcher) is not only to send or not send a warning, but to configure and send differentiated products for many different areas (relates e.g. to principle 1). Regarding the operators’ goals there might be a need to enlarge the warning product configuration section for guidance. Avoid excessive menuing and windowing in order to highlight important information (principle 4). Indicate the presence of prototypical situations (if known – principle 6), indicate sensors whose information is missing within the map view (9), support sensor reliability assessment (10) e.g. using luminance levels, and consider presenting likeliness of errors (e.g. 22 %) together with reliability (e.g. 78%). Clearly distinguish between known data and inferred data (14). We hope this exemplifies a method in practice.

Measurements of situation awareness may be explicit or implicit, objective and subjective. With explicit measures at certain intervals the task or simulation is temporarily frozen and subjects are presented a set of predetermined multiple-choice questions about the situation or “real-time probes” embedding open questions as verbal communications during the task. Explicit measures are those which seek to capture how people actually perceive and understand the key elements of the situation. They involve the use of “probes” or questions designed to prompt subjects to self-report their actual SA. Probe techniques include the use of open questions embedded as verbal communications during the task (known as “real-time probes”). This method is less intrusive and more “naturalistic” than artificially interrupting and freezing the task [6].

Implicit measures are those in which the state of someone's SA is inferred from indirect (but objective) evidence, such as task performance analysis, communications analysis or, physiological data. Subjective measures measure perceived quality of SA through self-ratings, observer-ratings, or peer-ratings to evaluate Situation Awareness in teams. The general recommendation for measuring SA is that, when possible, several measures of SA should be utilized to ensure concurrent validity [5] and to provide a balanced, informative assessment.

For our evaluation we only had the three static screens as “frozen” probes to work with. Besides the heuristic evaluation we conducted usability tests with 10 users applying eyetracking analysis as implicit measures and self-reports about the understanding of the simulated situation as explicit measure.

4. Heuristic Evaluation

Within the heuristic evaluation two independent experts inspected each of the four screens separately and then at the interaction between them. Expert inspection was the first step of evaluation of the static prototype consisting of three screens. We analyzed the conformance of the current interface design with the initially described guidelines.

Methodologically the expert inspection starts with usage scenarios (respectively typical user workflows). We analyzed, in how far the current interface concepts and design conform to the safety-critical principles supporting expert users. For each potential issue the problem or idea was listed with a reference to a usage scenario, a dialogue principle (DiaP 1-7) or an SA-principle (8-19) and where appropriate one of the perspectives (situation=SP, observation=OP, decision=DP).

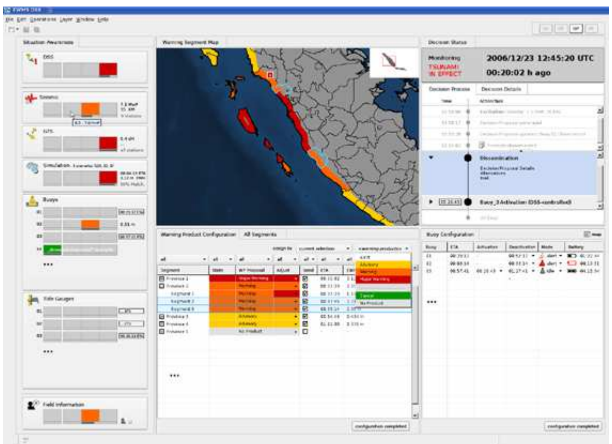

Image 3: The decision perspective of the system.

The heuristic analysis yielded findings like a missing association of the provinces in the map and in the table (for the decision perspective, dialogue principle 1). This stimulated the discussion if the integration of names in map could be a viable solution.

Conducting heuristic evaluation potential problems can only be estimated since expert evaluation works without user involvement. In order to integrate real user feedback complementary user tests with eye-tracking have been conducted.

5. Eye-tracking Evaluation

The duration for a single test was about 45 minutes. Test tasks have been derived from questions by DLR related to 9 use cases. Fixation length served as an indicator of specific levels of information processing.

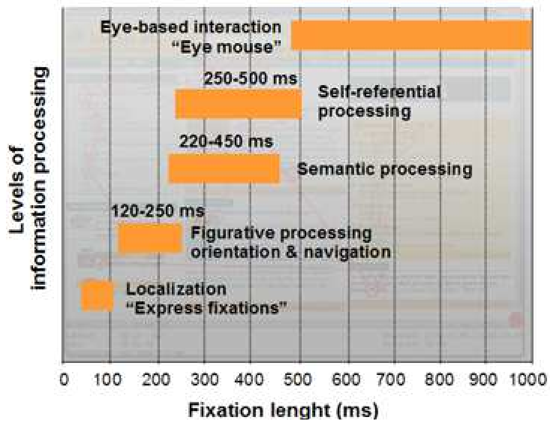

Image 4: Relation of processing and fixation times.

Focussing on tasks within the individual screens we tested with the Tobii Eyetracking system. The Tobii T60 Eye Tracker is integrated into a 17” TFT monitor. It is ideal for all forms of eye tracking studies with stimuli that can be presented on a screen.

Focussing on tasks that involve orientation between screens we test with the head-mounted systems by SMI. The iView X™ HED is the latest generation mobile eye-tracking system, combining full freedom of movement with easy setup and efficient operation. Robust and reliable, the use of a lightweight headset and tablet PC makes it suitable for indoor and outdoor use.

With the head-mounted system the following tasks were given to the user:

- Please first gain an initial overview of the three screens. What do you notice first?
- Which signals would you expect when there is new information?
- Gain an overview of the individual sensor data. Can you estimate the potential danger of a Tsunami (anticipation)?

- How high is the probability that a measured value is correct (comprehension)?
- Which value on the Richter scale has the earthquake that caused the Tsunami?
- Where do you expect to find the details or explanation of the indicators leadings to the Major Warning?
- Create and send a warning.
- How do you judge the Major Warning in the “Decision View”? Judge the situation using the information show to you on all three screens.

Image 5: User testing with the head-mounted system.

Thinking aloud was allowed but not enforced. Eyetracking data has been recorded and analysed using gaze-point and hot-spot analysis. The following image shows as an example the merged images of four test participants gaining a first impression of the decision view with the Tobii Eyetracking system (red areas indicating highest initial attention).

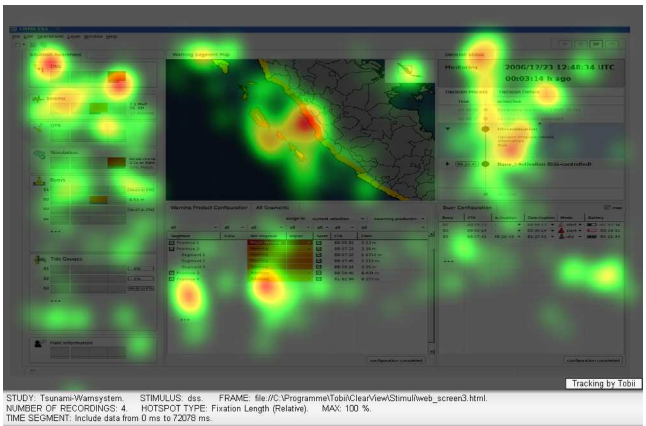

Image 6: Eyetracking results from the Tobii eyetracker.

After each test users answered standardized questionnaires relating to the tasks and design elements on the screens. All issues regarding general usability and regarding the perception, comprehension and anticipation of the situation that have been observed during the test or articulated by the test persons were documented.

Recurring issues from the heuristic inspection and the empirical user-testing were clustered together into topics and prioritized with respect to their potential to obstruct situation awareness on one of the three levels and to result in false (positive or negative) decisions.

Addressing all of the identified issues and topics, also reviewing the guidelines recommendations for redesign have been derived in order to improve the spatial arrangement and grouping of notifications on the displays, color and symbol coding, compatibility and expectancy.

6. Conclusion

We discussed challenges in user-centered design of safety-critical applications and presented the basic concept of an early warning system. Extending dialogue principles with situation awareness heuristic evaluation pointed out critical issues and triggered discussion. Their completeness and validity, or applicability to other domains is questionable, maybe left to a meta-study on research-based evidence for guidelines to design safety-critical applications and empirical in-depth evaluation of each single guideline as it has been conducted in [7]. Trying to approach such control situations we reviewed literature and conducted eye-tracking explorations with a head-mounted system and a single screen tracker. Applying this mix of methods including literature reviews, interviews, and usability testing with eyetracking observation allowed identifying and solving potential perceptual, cognitive and prognostic questions and improving situation awareness and well-informed decision support.

Heuristic evaluation indicated potential usability problems of the GUI prototype due to missing consistency and user guidance through the subsequent steps, as well as inconsistent design between the perspectives may make changes on one perspective difficult to recognize. Eyetracking data analysis indicated additional potential problems.

Within our redesign proposals we addressed the mapping of maps and information, integration and alignment of field and data labels, and a consistent design of timelines. Examples of resulting recommendations that could be implemented include:

- Rearranging and resizing views / info boxes
- Integrating additional information and options
- Clarifying user guidance by increasing focus on fewer options (for instance in the lower middle part of the decisive perspective).

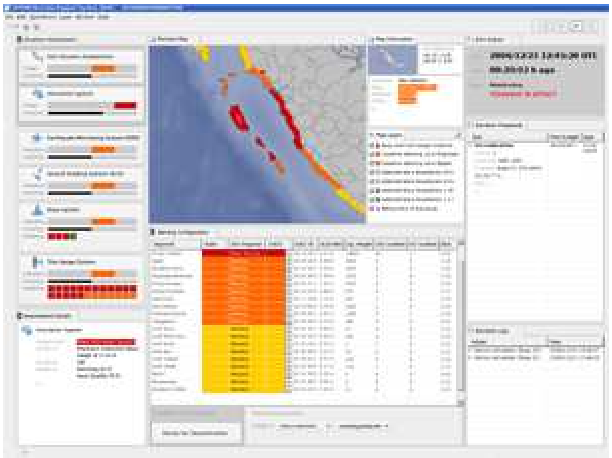

Image 7: Current version of the decision perspective

As valuable next steps paper prototyping exercises with design alternatives and eyetracking analysis with an interactive prototype have been recommended. Currently we evaluate a transfer of some principles and insight gain from the project to the realm of air-traffic control (slot management and flight planning).

7. Outlook

Based on the results of the DSS GUI evaluation described in this paper, the DSS GUI has been improved and extended.

On November 11, 2008, the Indonesian president, Susilo Bambang Yudhoyono, officially launched the Indonesian Tsunami Early Warning System (InaTEWS) at the Indonesian Agency for Meteorology, Climate and Geophysics (BMKG) in Jakarta. A test and commissioning phase will now ensure that the system can be optimized according to the needs of BMKG and is ready for operations. The System will not only serve for early warning purposes in Indonesia but is also planned to serve as a so-called Regional Tsunami Watch Provider (RTWP) Center for the entire Indian Ocean in tight cooperation with warning centres in other Indian Ocean rim countries. According to RTWP requirements and user feedback, the DSS GUI will be further improved in the future.

8. Acknowledgments

Research by Henning Breuer and Mitsuji Matsumoto is supported by NICT.

GITEWS is being developed by a number of scientists and engineers of the Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences (GFZ, Project Lead), the Alfred Wegener Institute for Polar and Marine Research (AWI), the German Aerospace Center (DLR), the GKSS Research Centre (GKSS), the German Marine Research Consortium (KDM), the Leibniz Institute of Marine Sciences (IFM-GEOMAR), the United Nations University Institute for Environment and Human Security (UNU), the Gesellschaft für Technische Zusammenarbeit (GTZ) and the Federal Institute for Geosciences and Natural Resources (BGR) as well as Indonesian and international partners. The project is funded by the German Federal Ministry for Education and Research (BMBF), Grant 03TSU01.

Literature

[1] Adam, E.C. (1993). Fighter cockpits of the future. Proceedings of the 12th DASC pp318-323. The 1993 IEEE/AIAA Digital Avionics Systems Conference.

[2] Endsley, M.R. (2000). Theoretical Underpinnings of Situation Awareness. A critical review. In M.R. Endsley / D.J. Garland. (Ed.). Situation Awareness Analysis and Measurement (pp. 3-32). Lawrence Erlbaum Associates. Mahwah, NJ.

[3] Endsley, M.R., Bolté, B. & Jones, D.G. (2003). Designing for Situation Awareness. An Approach to User-Centered Design. Taylor & Francis.

[4] DIN EN ISO 9241-110: Grundsätze der Dialoggestaltung, Beuth Verlag, Berlin 2006.

[5] Harwood, K., Barnett, B., & Wickens, C. D. (1988). "Situational Awareness: A Conceptual and Methodological Framework". Proc. of the 11th Symposium of Psychology in the Department of Defense.

[6] Jones, D.G., & Endsley, M.R. (2000). Can real-time probes provide a valid measure of situation awareness? In Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium. Savannah, GA: SA Technologies, Inc.

[7] Usability.gov (2006). Research-Based Web Design & Usability Guidelines. (Online: http://usability.gov/ pdfs/guidelines.html)