DR 4.4.4: Summative evaluation and theory of the setting-up and usage of adaptive HRI

Tina Mioch*, Mark A. Neerincx*, Valsamis Ntouskos†, Fiora Pirri†, Nanja J.J.M. Smets*, Joris Janssen*

*TNO, Kampweg 5, 3769 DE Soesterberg, The Netherlands †ROMA, Rome, Italy
⟨tina.mioch@tno.nl⟩

Project, project Id:	EU FP7 NIFTi / ICT-247870
Project start date:	Jan 1 2010 (48 months)
Due date of deliverable:	M48
Actual submission date:	December 2013
Lead partner:	TNO
Revision:	final
Dissemination level:	PU

WP4 aims at enhanced effectiveness and efficiency of joint human-robot exploration by dynamic task load distribution among rescuers and robots, preventing overload and optimizing attention allocation. In year 4, we focused on the integration of the prototype components into an integrated, operational system. The cognitive task load model and the team awareness display were integrated into the NIFTi system, providing the (adaptive) working agreement policies, personalized views on the current (shared) knowledge-base (i.e. supporting situation and team awareness), and context-sensitive information exchanges (e.g., "smart questions"). In addition, 3D eye-tracking experiments were conducted to further develop computational visual attention models aimed at solving top-down search tasks.

We applied the methodology for *situated Cognitive Engineering* (sCE) during the 4 years of the NIFTi project to establish a requirements baseline with a sound and practical design rationale. Over the years, NIFTi-robots' level of autonomy, user model, and team membership were enhanced for disaster management scenarios with increasing scale and complexity (from

tunnel and train accident to earth quake). The current requirements baseline, with its design rationale, provides a sound basis for (1) implementing state-of-the-art collaborative rescue robots, and (2) further developing of such robots with higher levels of persistence. The last year, we extended the sCE-methodology by incorporating interaction design patterns that provide a structured format to capture and share design knowledge on the communication level (i.e. the shape of the interaction).

Executive Summary

In year 4, WP4 refined the user and attention models to improve the joint human-robot exploration. Core support functions, with their design rationale, have been developed to enhance shared situation awareness and to harmonize the task load allocation. These functions concern both the accommodation of adaptable work organizations and the adaptations of the Human-Robot Interactions (HRI) to different operational contexts. Their iterative and incremental development has been guided by the NIFTi roadmap, focusing on human-robot cooperation in realistic complex settings. These settings encompass a distributed human-robot team – with both command post and in-field operators – in a chaotic earthquake terrain, implying high demands for coordination and communication.

The support functions were further refined and integrated into the overall NIFTi prototype, studying and demonstrating more advanced adaptive support. We improved the Cognitive task Load and Emotional State models and integrated them into the final prototype. Furthermore, they were incorporated into a formal framework for dynamic task allocation and adaptive dialogues. Methods for balancing the information transfers were refined (i.e., the policies for information exchange of team members with specific roles and capabilities). In the last year, we extended the functional design with re-usable shape specifications, i.e., we provided, implemented and evaluated interaction design patterns for enhanced scheduling and balancing of tasks (via the team awareness display) and for supporting the communication in the team. A data-set was collected for conducting 3D eye-tracking experiments to further develop computational visual attention models for top-down search tasks.

Important results of these research efforts on adaptive support were the following. First, *refinement and integration* of (1) the user model (for cognitive and affective load) for *dynamic load balancing*, and (2) the context-sensitive situation and team awareness display for sharing information and establishing the Right Message at the Right Moment in the Right Modality, or $(RM)^3$, *balancing the communication.* Second, *extension* of the prototype design with working agreement policies and interaction design patterns to establish coherent and re-usable specifications of human-automation interaction at the communication level. Third, statistically reliable eye-tracking *data collection* of search patterns with a concrete task to evaluate an attention model.

Role of Cognitive Task Load and Selectional Attention in NIFTi

NIFTi investigates cognitive architectures which can meaningfully sense, act, and cooperate with humans in real-life settings. WP4 contributes to this by building up knowledge on the user and supporting the user to stay in a continual workflow, by attuning the information processing and sharing to the task at hand. The three objectives of WP4 are contributing to this: (A) Cognitive taskload-based mechanisms for scheduling and balancing tasks, (B) Mechanisms for selectional attention based on mission briefs (presence of people, dangerous items) and context and (C) Mechanisms for determining the right time and right form of multi-modal dialogue based on situated context, user models (workflow), and cognitive task load. WP4 contributes to three project Objectives: Objective 2 Situated cognitive user models, Objective 3 User-adaptive human-robot communication and Objective 4 Flexible planning & execution.

Contribution to the NIFTi scenarios and prototypes

WP4 contributes to the scenarios and prototypes by including sections in the scenario where we expect to have high task load when using a less autonomous robot than the envisioned NIFTi robot, and which include areas the robot should allocate attention to (i.e., locations where victims might be).

In addition, in WP4, situations are identified, in which the scheduling and allocation of tasks and the balancing of information ensure a higher team performance. In year 4, the scenario has been extended to a complex earthquake terrain, in which cooperation and communication play an even more important role than the previous years. In WP4, it is researched how these team aspects can best be supported.

Furthermore, WP4 contributes to the prototype by making use of human-like search patterns for the robot, so that the human can relatively easily recognize objects that the robot "misses" (i.e., which are not recognized automatically). In addition, the behavior of the robot is easier to interpret for the user, because the search patterns displayed by the robot are similar to what the human would do and therefore expects.

1 Tasks, objectives, results

1.1 Planned work

The aim of WP4 is to improve the effectiveness and efficiency of joint exploration by dynamic task load distribution among human rescuers and robots, preventing overload and optimizing attention allocation.

In the reporting period for this deliverable, WP4 focused on scheduling and allocation of tasks based on task load (T4.1), selectional attention (T4.2), and cooperation between the team members, both robotic and human (T4.3).

By the work reported in this deliverable we have achieved the year 4 Milestone MS4.4:

MS4.4 - Working agreements for cooperation strategies. The objective is to develop methods to set the working agreements for dynamic task allocation and adaptive dialogues.

Scheduling and allocation of tasks T4.1 focuses on adapting the scheduling and allocation of tasks to the user, based on the capabilities and availability of the user and the robot. In this reporting period, the focus lies on evaluating and improving the cognitive task load model.

Selectional attention In this reporting period, the focus lies on collecting statistically reliable eye-tracking data in free-viewing mode in real cluttered environments, and on recording search patterns with a concrete search task in mind. This data is used to evaluate an attentional model.

Cooperation between team members T4.3 focuses on adapting the scheduling and balancing of the communication for cooperation. In this reporting period, the focus lies on balancing the information that is transmitted based on the role and capabilities of the person, and on working agreements.

1.2 Addressing reviewer's comments

The WP4 recommendations for the year 4 efforts made by the reviewers were the following:

- Overall recommendation: Focus on integration and integrating as many of the sub-systems that are not yet integrated but have been tested in the operational system.
- It is quite difficult to develop tools concerning cognitive load, to assess those tools under controlled conditions that also elicit stress from users

and to integrate such tools. This workpackage has made good progress in all of these areas. The review panel encourages the workpackage team to pursue this line of research.

We have taken these recommendations into account by firstly focusing on the integration of the developed systems into an integrated, operational system. The cognitive task load model has been integrated into the team awareness display, showing dynamically the current task load of the UGV operator to the other team members. In addition, the cognitive task load model has used as basis for the development of a dynamic task allocation model.

Furthermore, as is described below, further studies on workload assessment and balancing information based on the roles and capabilities have been conducted.

1.3 Actual work performed

1.3.1 Scheduling and allocation of tasks

T4.1 focuses on adapting the scheduling and allocation of tasks . In the following, the work is described that has been done in on this topic in reporting period 4.

Cognitive task load As described in DR4.4.3, based on [28], a cognitive task load model has been developed for real-time monitoring and, subsequently, balancing of workload on three factors that affect operator performance and mental effort: mental occupancy, level of information processing, and number of task set switches [6].

This cognitive task load model has been adapted to be able to handle more detailed and extended input of the task environment and was used as basis for a task allocation model (see DR4.4.3 and [16]).

During the end-user evaluation of year 4, for both models [6, 16], it was evaluated how the interpretation of the cognitive task load corresponds to the real workload of the fire fighters [10]. Throughout the mission, the UGV operator reported on his workload every two minutes. He did this by rating it on a scale from 1 to 5, where 1 was a very low workload and 5 a very high workload. We also measured his heart rate variability, galvanic skin response and facial expressions.

To assess model accuracy, the results from the self-reported workload and the dynamic workload models were plotted per participant (for the plots, please see [10]). It can be concluded that the adapted model has better results than the not adapted model, and that it's workload outputs (relatively) closely follow the users' subjective estimation of their cognitive state.

It was also evaluated whether there was a correlation between the cognitive workload of the operator (reported and derived by the model) and the time it took to respond to the two-minute signal to report the workload. However, no correlation was found.

1.3.2 Selectional Attention

In year 4, we refined the attention models to address the requirements for adequate, "natural", attention allocation in the complex earth quake scenario. There is a lack of reliable data-sets for search task in such environments. Therefore, our work focused on two tasks. At first, we collected statistically reliable eye-tracking data in free-viewing mode in real cluttered environments using the Gaze Machine. In order to keep participants motivated the goal was to remember as many details about the scene as possible during a limited period of time. Eye-tracking data were obtained by creating hit maps overlaid with 3D reconstructed scenes.

As a second task, we recorded search patterns with a concrete search task in mind, such as counting different instances of the same object. This was achieved in the same way as the first task. However, with such a dataset it becomes now possible to study human behavior in terms of evaluation of directed top-down visual attention.

As a pilot study we created an experimental setup with 3 participants to check the feasibility of our methodological approach. We are currently analysing the data and therefore present more detailed results at the end-review.

Preliminary results show that attentional models designed to work on static images or videos fail to predict fixations when they are applied in the real world environment. It can potentially lead to designing an attentional system comparable to human performance, when depth perception and space orientation are taken into account. Our future work will include the design of a general protocol to perform more experiments on visual search and algorithms to process eye-tracking data [29]. More details can be found in Annex §2.5.

1.3.3 Cooperation

T4.3 focuses on adapting the scheduling and allocation of communication to the user. In this section, we describe results regarding this topic for the reporting period 4.

An ontology for situation awareness for professionals in the high risk domain An earthquake has damaged several buildings. An Urban Search and Rescue (USAR) team reaches the area. The USAR team operates in a high risk domain with high task demands, danger (gas leaks), changing circumstances and a changing environment (further collapses can occur). The USAR team is searching for victims in the debris, but also building up knowledge about what happened, where the unstable and unsafe areas are, thus building up awareness about the situation. The USAR team in this scenario consists of humans and ground robots. All of the actors have their own capabilities, tasks, goals and their own awareness about the situation. Situation Awareness (SA) means that a person is able to perceive and comprehend what is happening in the environment and what the situation is, but also able to predict what happens when an event occurs [12]. The USAR personnel also has to know whether the information or awareness they have of the environment is still up to date. Because the environment they operate in is a dynamic environment, changing over time, their awareness has to be updated constantly [8]. Situation awareness is a complex concept, Endsley has provided a first, well-cited, definition of SA for certain tasks (e.g., pilots in a cockpit). However, there is criticism on the SA concept in literature. Questions are raised on how to measure SA, but also on how to distinguish SA from long-term memory? There are also researchers who question the SA concept altogether [46]. In our view, Endsley's concept of SA serves as a good basis (e.g., distinguishing perception, comprehension and projection), but it should be extended to incorporate the work demands and constraints of USAR operations for our research.

We have been working on the SA concepts relevant to professionals operating in the high risk domain for certain use cases. These SA concepts are documented in an ontology. The goal of this ontology is to define relevant concepts and relations regarding SA for our use cases. For the mindmap of the concept of situation awareness, please see Annex §2.6. This is used as a basis for the development of an ontology for SA.

Communication During the year 4 evaluation, five teams performed an USAR scenario at the test site in Prato, Italy. The communication between the teams was logged. Focus of the analysis was on the communication used by the team members (robots and humans). First the objects in the team awareness system were mapped on the ground truth data. Second, for each scenario, the main events logged in the system were plotted on a timeline (e.g., discovery of victims by an actor). Using the data, we defined three representative events. For these specific events, we looked in depth at the interaction and temporal aspect for all communication modalities: Radio communication, face to face communication, and communication via the team awareness display [23].

We identified forms of communication (e.g., orders, instruction, situation-reports). Looking at the forms, we defined an optimal modality to perform this form of communication. Using this perspective, we analyzed the data and compared the actual communication modalities with the defined opti-

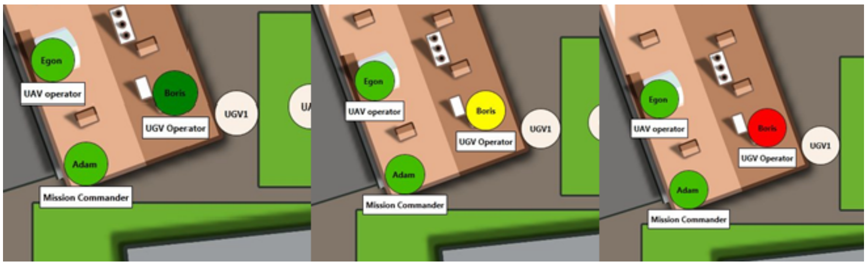

Figure 1: Representation of the operator's workload on the team awareness display. Green indicates a low workload, yellow a medium workload, red a high workload.

mum.

Combining the results, we propose five design patterns to improve the communication in general and team team awareness display specifically within USAR scenarios. For a detailed description of the design patterns, see Annex §2.4.

Task load of operator By implementing the CTL-model [6] and dynamically updating the status on the team awareness display, information sharing can be dynamically attuned to the operator's state. Every second, the CTL-module updates a diagnosis of the operator's cognitive state, which then is displayed on the team awareness display, see Figure 1. This makes it possible that team members and the mission commander have insight in the workload of the operator and can pro-actively re-allocate tasks based on this insight.

During the end-user evaluation of year 4, a first evaluation of representing the operator's cognitive state was done [10]. It was evaluated whether team members pay attention to and adjust their behaviour based on the other member's cognitive state. This was evaluated through a questionnaire, including Likert-scale (1-5) questions in addition to open questions. Also, after the experiment, discussion sessions were held in which fire fighters could give feedback on the interface. The answers by the fire fighters indicate that they did not notice the change(s) on the interface of the operator's cognitive state. However, the importance of knowing other team members' cognitive state was pointed out by fire fighters during the discussion sessions after the experiment. This indicates that team members would adjust their behaviour based on the other team members' cognitive state, and that the current representation of the operator's cognitive state on the interface is not sufficiently clear and present and needs to be improved.

Interaction Design Patterns for Coherent and Re-usable Shape Specifications of Human-Robot Collaboration. We extended the methodology for situated Cognitive Engineering (sCE) by incorporating interaction design patterns into the methodology. So far, the sCE methodology provided a specification of the (functional) user requirements with the related scenarios, use cases and claims (i.e. a specification at the task level). Interaction design patterns provide a structured format to capture and share design knowledge at the communication level (i.e. the shape of the interaction).

The extended methodology has been applied to the development of human-robot cooperation in the urban search and rescue domain, or more specifically to team-awareness display functionalities, see Annex §2.3. We found that a design specification can be valid on a task level, while the evaluation shows sub-optimal results because of a moderate communication level. Based on this evaluation result a design improvement on the communication level has been proposed without the need to adjust the task level design solution [27].

1.4 Relation to user-centric design

As work and user modeling are core of this workpackage, user-centric design has been applied from the start. The users were involved in the situated support needs analysis, both during the development of situated scenarios as in the evaluation. Furthermore, humans were the inspiration for the selectional attention mechanism and the validation source (does the robot do the same as the human). By extending the current scenario and claims analysis approach [32] towards a more formal requirements' justification in which claims incorporate user theories, a "situated" user theory was formulated and tested with the end-users.

1.5 Relation to the state-of-the-art

Technical progress of robots for search and rescue is going on, and there is a growing body of relevant research, e.g., the Human-Agent Robot Teamwork (HART) workshops. However theoretical and empirical foundations are lacking for real-world design proposals for integrated, context-sensitive cognitive systems.

Workload First, current practices show that workload is still a major bottleneck in human-robot interaction. Robot operators spend up to 60% of their time attempting to establish and maintain situation awareness, and can cause "human"errors, such as driving mistakes and misses of victims [18], [19], [17]. We refined and implemented models of Cognitive Task Load (CTL) and Emotional State (ES) for adapting the individual (often mobile) interface to the workload and present the workload distribution to the team [5]. In this way, we could, real-time, analyse workload distributions in complex and realistic disaster response settings automatically.

Situation and team awareness Second, situation and team awareness prove to show shortcomings during diaster responses and sound support systems are needed that **efficiently** acquire, assess and apply actual information in time [31, 40, 9]. We extended current concepts of *Situation Awareness (SA)* (e.g., [11], [12], [46]) to represent humans and robots as part of each other's environment (i.e., the meaning is depending on the "owner(s)" and operational objectives (cf. [1], [46], [13])). Based on the SA-model, information exchange is optimized by presenting it in aggregated views, at higher levels of abstraction [33] (incl. content-based, context-sensitive filtering [38]). By integrating these models, real-time adaptive robot behavior during sense-making is supported [44]. Not only tactical information which is known is shared, but also tactical information which is known to be unknown is shared [37], [7]. For example, an area which has not been explored in a prior sortie should be marked as such for the next team. However, even if the area has been explored in a prior sortie, the absence of a fact X does not necessarily mean that fact X is not the case. This can be because the prior sortie was directed at finding out other information than fact X.

Selectional attention Third, eye-tracking experiments are important for providing ground truth data necessary to evaluate computational visual attention models. There are two types of visual attention: bottom-up and top-down attention. Bottom-up attention, is driven by object properties. Top-down attention, on the other hand, takes place when there is a specific search task at hand. While there exist several datasets for bottom-up attention only a small number of experiments have aimed in collecting eye-tracking for specific search tasks. As an example of such datasets, experiments for driving a car and other task-oriented activities, in a virtual environment (video games) were done in [2].

Most of the available datasets are based on two dimensions using static images or video sequences. Depth perception though, has been shown to have an important role in human attention behaviour [45]. Based on this, some attempts have been done to collect eye-tracking data either from free-viewing stereo images, like in [24], or in artificially created [43] and real 3D environments [34].

This work aims at producing a dataset with 3D eye-tracking data from several participants for creating a benchmark for saliency algorithms and computational attention models aimed at solving top-down search tasks.

1.6 Overall WP4 progress

Over the last four years of the NIFTi project, WP4 has made a set of significant contributions to the NIFTi project. We applied the methodology for *situated Cognitive Engineering* (sCE) to establish a requirements baseline with a sound and practical design rationale. Over the years, NIFTi-robots' level of autonomy, user model, and team membership were enhanced for disaster management scenarios with increasing scale and complexity (from tunnel and train accident to earth quake). The current requirements baseline, with its design rationale, provides a sound basis for (1) implementing state-of-the-art collaborative rescue robots, and (2) further developing of such robots with higher levels of persistence. In the following, the main contribution of different topics is shortly described for each task.

1.6.1 Scheduling and allocation of tasks

T4.1 focused on adapting the scheduling and allocation of tasks to the user.

Cognitive task load and emotions A cognitive task load model was designed, implemented and tested to keep track of an operator's mental workload in real time, both quantitatively (amount of workload) and qualitatively (cognitive state) [4, 5].

This model is based on three factors that affect operator performance and mental effort: mental occupancy, level of information processing, and number of task set switches [3]. This model was evaluated in a user study, in which five participants drove the USAR robot in over 16 hours of driving time [6]. In addition, it was evaluated how the interpretation of the cognitive task load corresponds to the real workload of the fire fighters (e.g., year 4 [10]). The model has been integrated into the team awareness display; on the display, the current cognitive state of the operator is displayed. We evaluated presenting this information. It can be concluded that the fire fighters believed this information to be very important, but that the current representation on the display is not ideal and needs to be improved.

In addition, we researched whether the emotional state of the robot operator is measurable to be able to interpret the actual emotional state to adapt the task distribution between the different task members. As a first step, we evaluated whether it is possible to invoke a realistic emotional state in a scenario-based evaluation [15], and how we can measure the emotional state of the robot operator in a realistic environment, in dynamic and complex tasks [36, 15, 26].

Task allocation We introduced a high-level framework for dynamic task allocation. The framework details how context information can be used to find possible role assignments for actors and to evaluate these role assignments. It also describes the important concepts in context information that influence team performance and can be used to dynamically allocate tasks [14].

The framework was used as a basis for designing a model for adaptive automation, taking into account the cognitive task load of the operator and the coordination costs of switching to a new task allocation. Based on these two context factors, it finds the optimal level of autonomy of a robot, separately for all tasks that need to be executed. We evaluated this model in a small experiment [16].

1.6.2 Scheduling and allocation of communication

T4.3 focused on adapting the scheduling and allocation of tasks to the user.

Situation awareness We have been working on the SA concepts relevant to professionals operating in the high risk domain for certain use cases. These situation awareness concepts are documented in an ontology. The goal of this ontology is to define relevant concepts and relations regarding situation awareness for our use cases.

Cooperation To attune the level of robot autonomy to situational demands, we studied the constraints of robot's team-membership. A first experiment provided requirements for the communication and task load that should drive the actual setting of robot autonomy and, consequently, the setting of working agreements [21, 20, 22].

Team awareness display Adapting the display to the momentary user needs and context. An agent-based architecture was developed that enables real-time decisions of (adaptive)information presentation (i.e., to establish context-sensitive "Right Messages at the Right Moment in the Right Modality" (called situated $(RM)^3$) [42]. If, how and when information is delivered to the user is made dependent on the user's state, task, context, and available devices and services. This work has been done in close cooperation with tasks in WP3. Each year (from year 2 onwards), a new version of this team awareness display was evaluated and refined [30, 41, 25].

In addition, a set of extendable OWL ontologies was specified, and corresponding core $(RM)^3$ functions were implemented and tested (information push and pull, user adaptation, multi-device, and policy management) [39].

Situated Cognitive Engineering We applied the methodology for situated cognitive engineering during the 4 years of the NIFTi project. The methodology consists of three components: the operational, human factors, and technological analyses to establish a sound and practical design rationale, the specification and maintenance of the requirements baseline, and the evaluation by means of simulation or a prototype, to validate and refine the requirements baseline.

The methodology entails an iterative and incremental research & development process. Over the years, NIFTi-robots' level of autonomy, user model, and team membership enhanced for disaster management scenarios with increasing scale and complexity (from tunnel and train accident to earth quake). The current requirements baseline, with its design rationale, provides a sound basis for (1) implementing state-of-the-art collaborative rescue robots, and (2) further developing of such robots with higher levels of persistence.

Each year, evaluations have been done with end-users, which led to a refinement of the requirements baseline (see WP7). We extended the methodology by incorporating interaction design patterns into the methodology [27]. Interaction design patterns provide a structured format to capture and share design knowledge, in our methodology focusing on the communication level (i.e. the shape of the interaction). The extended methodology has been applied to the development of the team awareness display. Based on this evaluation, design improvements could be proposed on the communication level without the need to adjust the task level design solution (i.e., the use cases, user requirements and claims).

2 Annexes

2.1 Giele et al. "Dynamic Task Allocation for Human-Robot Teams"

Bibliography Tinka R. A. Giele, Tina Mioch, Mark A. Neerincx, and John-Jules Meyer. Dynamic Task Allocation for Human-Robot Teams. *Under review.* [16]

Abstract As technology advances, artificial agents such as robots are increasingly deployed in tasks in complex and dynamic environments. More and more, these sophisticated robots will work together with human agents in a team, leading to the need for research into cooperation in mixed human-robot teams. An important aspect of cooperation is dynamic task allocation.

In this paper, a framework for dynamic task allocation, aimed at improving team performance in mixed human-robot teams is presented. The framework specifies how context information is used to find the optimal task allocation for a team. One important context factor is the cognitive task load (CTL) of a human agent. Based on the framework, a model for adaptive automation is designed that takes the CTL of an operator and the coordination costs of switching to a new task allocation into account. Based on these two context factors, it finds the optimal level of autonomy of a robot for each task. The model is instantiated for a single human agent cooperating with a single robot in the urban search and rescue domain. A first experiment is conducted aimed at testing the model. Some encouraging results are found: the cognitive task load of participants mostly reacted to the model as intended. Furthermore, important focus points for improving the model are identified such as taking into account more context information, e.g. capabilities and preferences.

Relation to work performed This work directly contributes to T4.1.

2.2 Eilers "Robot for Urban Search and Rescue – Modelling Cognitive Task Load"

Bibliography Renate Eilers. Robot for Urban Search and Rescue – Modelling Cognitive Task Load. Research project, Artificial Intelligence, Faculty of Humanities, Utrecht University, October 2013. [10]

Abstract In this document, a model based on the cognitive task load model (CTL model)developed by Neerincx [28] is evaluated. This model estimates the cognitive task load for the UGV-operator in real time, based on the tasks this operator is executing [5]. Two implementations of this model were tested: the Automatic cognitive task load model (ACTL) [3] and the Task reallocation model (TR) [14]. We also investigate what the effects of explicitly showing a team members current task load are on the other team members behaviour. It can be concluded that the TR model has better results than the ACTL model, and that it's workload outputs (relatively) closely follow the users' subjective estimation of their cognitive state. In addition, results indicate that the fire fighters did not notice the change(s) on the interface of the operator's cognitive state. However, the importance of knowing other team members' cognitive state was pointed out by fire fighters during the discussion sessions after the experiment.

Relation to work performed This work directly contributes to T4.1.

2.3 Mioch et al. "Interaction Design Patterns for Coherent and Re-usable Shape Specifications of HumanRobot Collaboration "

Bibliography Tina Mioch, Wietse Ledegang, Rosie Paulussen, Mark A. Neerincx, Jurriaan van Diggelen. Interaction Design Patterns for Coherent and Re-usable Shape Specifications of HumanRobot Collaboration. *Under review.* [27]

Abstract Sharing and re-using design knowledge is a challenge for the diverse multi-disciplinary research and development teams that work on complex and highly automated systems. For this purpose, a situated Cognitive Engineering (sCE) methodology was proposed that specifies and assesses the functional and shape requirements with their design rationale in a coherent and concise way. This paper presents this approach for the development of human-robot collaboration, focusing on the application of interaction design patterns to capture and share design knowledge on the shape of the human-robot interaction (i.e., the communication level). The sCE case study in the urban search and rescue domain provided the specification and assessment of functions and shape of a team-awareness display. Twenty fire fighters participated as operator of a ground or aerial robot, in several realistic earth quake scenarios to assess the functions and shapes of this display in different settings. It showed that the functions (i.e., the task level requirements and rationale) were valid, while the shape (communcation level) was (yet) sub-optimal. Based on this evaluation result, a design improvement on the communication level has been proposed without the need to adjust the task-level design solution.

Relation to work performed This work directly contributes to T4.3.

2.4 Janssen et al. "Towards Shared Awareness for a Human-Robot Team "

Bibliography Joris B. Janssen, Jouke Rypkema, Bas Holleman. Towards Shared Awareness for a Human-Robot Team. Technical report, TNO, December 2013. [23]

Abstract During the year 4 evaluation, five teams performed an USAR scenario at the test site in Italy. The communication within the teams was logged. Focus of the analysis was on the communication used by the team members (robots and humans). First the objects in the team awareness system were mapped on the ground truth data. Second, for each scenario, the main events logged in the system were plotted on a timeline (e.g., discovery of victims by an actor). Using the data, we defined three representative events. For these specific events, we looked in depth at the interaction and temporal aspect for all communication modalities: Radio communication, Face to face communication, communication via the team awareness display.

We identified forms of communication (e.g., orders, instruction, situation-reports). Looking at the forms, we defined an optimal modality to perform this form of communication. Using this perspective, we analyzed the data and compared the actual communication modalities with the defined optimum. Combining the results, we propose five design patterns to improve the communication in general and the team awareness display specifically within USAR scenarios.

Relation to work performed This work directly contributes to T4.3.

2.5 Ekaterina Potapova, Valsamis Ntouskos, Astrid Weiss, Michael Zillich, Markus Vincze, Fiora Pirri, "A Pilot Study on Eye-tracking in 3D Search Tasks"

Bibliography Ekaterina Potapova, Valsamis Ntouskos, Astrid Weiss, Michael Zillich, Markus Vincze, Fiora Pirri, "A Pilot Study on Eye-tracking in 3D Search Tasks" [29]

Abstract Eye-tracking is an important step in evaluation of computational visual attentional models in terms of comparison to human visual perception. Nowadays, there exist lots of free-viewing datasets. However, not so many experiments were made to collect eye-tracking for specific search tasks. Search tasks is a subclass of top-down visual attention. Experiments that aim to investigate top-down visual attention collect eye-tracking data with a specific task in mind, such as driving a car, for example [1]. However, collecting eye-tracking data for pure visual search tasks still remains an unsolved and challenging problem. Such datasets can be used as benchmark for saliency algorithms and computational attention models aimed to solve top-down search tasks. In our paper we focus on two tasks. At first, we collect statistically reliable eye-tracking data in free-viewing mode in real cluttered environments using the Gaze Machine. Cluttered environment was aiming to represent a small part of a kids room and consisted of a set of different toys. In order to keep participants motivated the goal was to remember as many details about the scene as possible during a limited period of time. Eye-tracking data is obtained by creating hit maps overlaid with 3D reconstructed scenes. As a second task, we recorded search patterns with a concrete search task in mind, such as counting different instances of the same object. This was achieved in the same way as the first task. However, with such a dataset it becomes now possible to study human behavior in terms of evaluation of directed top-down visual attention.

Relation to WP In this work we addressed the issue of creating a benchmark for visual attention models based on eye-tracking data in realistic 3D environments. As an example a setup with cluttered toys was created. Participants performed two different tasks: free-viewing and counting. Pilot study showed that classical contrast-based models fail to predict human fixations in real 3D environments. Next steps will be to create a unified 3D reconstructed model of the scene and map fixation points onto this model. Our future work will also include the design of a general protocol to perform more experiments on visual search and algorithms to process eye-tracking data.

2.6 Smets "A Mindmap of Situation Awareness as Preparation for an Ontology in the USAR Domain "

Bibliography Nanja J.J.M. Smets. A Mindmap of Situation Awareness as Preparation for an Ontology in the USAR Domain. TNO, December 2013. [35]

Abstract This document presents a mindmap of the concept of situation awareness. This is used as a basis for the development of an ontology for situation awareness.

Relation to work performed This work directly contributes to T4.3.

References

[1] A. Blandford and W. Wong. Situation awareness in emergency medical dispatch. *International Journal of Human-Computer Studies*, 2004.

[2] A. Borji, D. Sihite, and L. Itti. Probabilistic learning of task-specific visual attention. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 470–477, 2012.

[3] T. R. Colin. Modeling the cognitive state of urban search and rescue robot operators in real time. Master's thesis, Cognitive Artificial Intelligence, Faculty of Humanities, Utrecht University, February 2013.

[4] T. R. Colin. Nifti – natural human-robot cooperation in dynamic environments. research project, Artificial Intelligence, Faculty of Humanities, Utrecht University, March 2012.

[5] T. R. Colin, T. Mioch, N. J. J. M. Smets, and M. A. Neerincx. Estimating an operator's cognitive state in real time: a user modeling approach. In *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication 2012 (Ro-Man 2012)*, Paris, France, September 2012.

[6] T. R. Colin, N. J. J. M. Smets, T. Mioch, and M. A. Neerincx. Real time modeling of the mental workload of an urban search and rescue robot operator. under review, 2014.

[7] T. de Greef, R. van der Kleij, L. Brons, W.-P. Brinkman, and M. Neerincx. Observability displays in multi-teams. In *10th International Conference on Naturalistic Decision Making (NDM)*, 2011.

[8] F. T. Durso, K. A. Rawson, and S. Girotto. Comprehension and situation awareness. *Handbook of applied cognition*, 2:163–193, 2007.

[9] T. E.Drabek and D. McEntire. Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature. *Disaster Prevention and Management*, 12(2):97–112, 2003.

[10] R. Eilers. Robot for urban search and rescue – modelling cognitive task load. research project, Artificial Intelligence, Faculty of Humanities, Utrecht University, October 2013.

[11] M. Endsley. Measurement of situation awareness in dynamic systems. *Human Factors*, 47:65–84, 1995.

[12] M. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 47:32–64, 1995.

[13] Y. Gatsoulis, G. Virk, and A. Dehgani-Sanij. On the measurement of situation awareness for effective human-robot interaction in teleoperated systems. *Journal of Cognitive Engineering and Decision Making*, 4:69–89, 2010.

[14] T. R. A. Giele. A framework for dynamic task allocation instantiated for cognitive task load-based adaptive automation. Master's thesis, Cognitive Artificial Intelligence, Faculty of Humanities, Utrecht University, 2013.

[15] T. R. A. Giele. Do firefighters show emotion in the urban search and rescue domain? research project, Cognitive Artificial Intelligence, Faculty of Humanities, Utrecht University, January 2013.

[16] T. R. A. Giele, T. Mioch, M. A. Neerincx, and J.-J. Meyer. Dynamic task allocation for human-robot teams. Under review, 2014.

[17] L. Gunawan. *Crowdsourced disaster response for effective mapping and wayfinding*. PhD-thesis Delft University of Technology, 2013.

[18] L. T. Gunawan, H. Alers, W. Brinkman, and M. A. Neerincx. Distributed collaborative situation-map making for disaster response. *Interacting with Computers*, 23(4):308–316, 2011.

[19] L. T. Gunawan, S. Fitrianie, Z. Yang, W. Brinkman, and M. A. Neerincx. Travelthrough: A participatory-based guidance system for traveling through disaster areas. In *Proceeding of Computer Human Interaction, CHI 2012*, pages 241–250, Austin, Texas, USA, May 2012.

[20] C. H. G. Horsch. Urban search and rescue robots: The influence of team membership of robots on team performance. Master's thesis, Eindhoven University of Technology, April 2012.

[21] C. H. G. Horsch, N. J. J. M. Smets, and R. H. Cuijpers. Exploring performance and situation awareness in usar unit tasks. research project, Eindhoven University of Technology, 2012.

[22] C. H. G. Horsch, N. J. J. M. Smets, M. A. Neerincx, and R. Cuijpers. Revealing unexpected effects of rescue robots' team-membership in a virtual environment. In T. Comes, F. Fiedrich, S. Fortier, and L. Geldermann, J. Yang, editors, *Proceedings of the 10th International ISCRAM Conference*, Baden-Baden, Germany, May 2013.

[23] J. B. Janssen, J. Rypkema, and B. Holleman. Towards shared awareness for a human-robot team. Technical report, TNO, December 2013.

[24] C. Lang, T. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan. Depth matters: Influence of depth cues on visual saliency. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, *Computer Vision ECCV 2012*, Lecture Notes in Computer Science, pages 101–115. Springer Berlin Heidelberg, 2012.

[25] J. Linssen. Using trex for usar-robot cooperation in nifti. Master's thesis, Utrecht University, 2011.

[26] T. Mioch, T. R. A. Giele, N. J. J. M. Smets, and M. A. Neerincx. Measuring emotions of robot operators in urban search and rescue missions. In *Proceedings of the 31st European Conference on Cognitive Ergonomics*. EACE, 2013.

[27] T. Mioch, W. Ledegang, R. Paulussen, and v. D. J. Neerincx, Mark A. Interaction design patterns for coherent and re-usable shape specifications of humanrobot collaboration. under review, 2014.

[28] M. A. Neerincx. Cognitive task load design: model, methods and examples. In *Handbook of Cognitive Task Design*, pages 283–305. Mahwah, NJ: Lawrence Erlbaum Associates, 2003.

[29] E. Potapova, V. Ntouskos, A. Weiss, M. Zillich, M. Vincze, and F. Pirri. A pilot study on eye-tracking in 3d search tasks. In *Proceedings of the First International Workshop on Solutions for Automatic Gaze-Data Analysis (SAGA)*, 2013.

[30] M. Pronk. Bidirectional trex. internship report, Haagse Hogeschool, July 2013.

[31] S. Ramaswamy, M. Rogers, A. Crockett, D. D. Feaker, and M. Carter. WHISPER - service integrated incident management system. *International Journal Of Intelligent Control And Systems*, 11(2):114–123, 2006.

[32] M. Rosson and J. Carroll. Scenario-based design. chapter 53, pages 1032–1050.

[33] M. Sam. Ecological approach to develop a shared mental model for collaboration at a distance during a crisis situation. *Master Thesis Media and Knowledge Engineering, TU Delft*, 2009.

[34] K. Santner, G. Fritz, L. Paletta, and H. Mayer. Visual recovery of saliency maps from human attention in 3d environments. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 4297–4303, 2013.

[35] N. J. Smets. A mindmap of situation awareness as preparation for an ontology in the usar domain. Unpublished working document, 2013.

[36] N. J. J. M. Smets, M. A. Neerincx, and R. Looije. Measuring user behavior in a complex usar team evaluation. In *Proceedings of the 8th edition of the International Conference on Methods and Techniques in Behavioral Research 2012*, Utrecht, The Netherlands, August 2012.

[37] N. J. J. M. Smets, J. W. Streefkerk, and M. A. Neerincx. Commanders dashboard: overview of tactical changes to improve situated decision making in the field. In *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics*, ECCE '10, pages 193–200, New York, NY, USA, 2010. ACM.

[38] J. Streefkerk, D. McCrickard, M. van Esch-Bussemakers, and M. Neerincx. Context-aware timing of notifications to manage interruption in a simulated surveillance task. *International Journal of Mobile Human Computer Interaction*, 4(3):1–27, 2012.

[39] J.-W. Streefkerk and A. Heuvelink. Smart questions context-dependent mobile information exchange for military operations. In *Proceedings of the 6th International Conference on Knowledge and Smart Technology (KST)*, 2014.

[40] K. B. C. Tierney and E. Kuligowski. Metaphors matter: Disaster myths, media frames, and their consequences in hurricane katrina. *The ANNALS of the American Academy of Political and Social Science*, 604:57–81, 2006.

[41] M. J. R. van den Heuvel. The effects of explicit unknowns on cooperation - an empiric study in urban search and rescue. Master's thesis, Cognitive Artificial Intelligence, Faculty of Humanities, Utrecht University, forthcoming.

[42] J. van Diggelen, M. Grootjen, E. M. Ubink, M. van Zomeren, and N. J. J. M. Smets. Content-based design and implementation of ambient intelligence applications. In *Proceedings of the 4th International Symposium on Ambient Intelligence 2013 (ISAmI 2013)*, Salamanca, Spain, 2013.

[43] D. Vishwanath and E. Kowler. Saccadic localization in the presence of cues to three-dimensional shape. *Journal of Vision*, 2004.

[44] K. E. Weick. *Sensemaking in organizations.* Sage Publications Inc., 1995.

[45] M. Wexler and N. Ouarti. Depth affects where we look. *Current Biology*, 18(23):1872 – 1876, 2008.

[46] C. D. Wickens. Situation awareness: Review of mica endsleyś 1995 articles on situation awareness theory and measurement. *Human Factors*, 50:397–403, 2008.