DR 7.1: Scenario-based evaluation in a large-scale static disaster area

Nanja Smets [†], Joachim de Greeff [*], Koen Hindriks[*], Maaike Harbers [*], Salvatore Corrao [‡], Norbert Pahlke [§], Guido van den Broek Humphreys [¶], Mario Gianni [‖], Valsamis Ntouskos [‖], Mark Neerincx [*] and the TRADR consortium

[*] *TUD, Delft, The Netherlands*
[†] *TNO, Soesterberg, The Netherlands*
[‡] *IFR, Italy*
[§] *FDDo, Dortmund, Germany*
[¶] *GB, Rotterdam, The Netherlands*
[‖] *ROMA, Rome, Italy*
⟨`nanja.smets@tno.nl`⟩

Project, project Id:	EU FP7 TRADR / ICT-60963
Project start date:	Nov 1 2013 (50 months)
Due date of deliverable:	December 2014
Actual submission date:	February 2015
Lead partner:	TUD
Revision:	final
Dissemination level:	PU

This document describes the scenario-based evaluation for multiple asynchronous sorties to assess a large-scale static disaster area. Several sessions with end-users are described including a qualitative session with end-users in an abandoned hospital in Calambrone (IT) that the Italian firemen use as a practice site.

Executive Summary

This report documents milestone MS7.1: *Scenario-based evaluation year 1* for WP7. This milestone consists of task T7.1: to perform a scenario-based analysis for multiple asynchronous sorties in a large-scale static disaster area. A **user needs analysis** was performed with the end-users by specifying the scenario that fulfils the goals of task 7.1 (multiple asynchronous sorties). The scenario describes an after-earthquake environment that needs to be explored (Figure 1). Based on the scenario, sessions with the whole consortium were held to **specify the socio-technical design rationale**. In the design rationale we specified requirements and linked them to the scenario.

The TRADR Joint Exercise (T-JEx) was held from 23 September until 2 October in Calambrone (Italy) at an abandoned military hospital site. T-JEx consisted of systems integration and testing, value assessment workshops and an **assessment with end-users** which constituted the qualitative study of the system; during this assessment the end-users performed an inspection of the hospital for chemical substances and structural damage using a team of people and robots (UGV and UAV).

Figure 1: Training site of Corpo Nazionale Vigili del Fuoco (VdF) in Calambrone, Italy, an old military hospital.

Role of scenario-based evaluation in TRADR

The TRADR-project takes a human-centred design approach (e.g., see ISO 9241-210:2010 "Human-centred design for interactive systems") with substantial end-user involvement from the start. In addition to more dedicated focus group, interview, questionnaire and observational studies, there is an "integrative evaluation" every year; first as formative evaluation and, subsequently, as summative evaluations. In the first year we defined the scenario with the end-user partners. The definition of the scenario and requirements specified in more detail how the system should operate within the context of an earthquake aftermath and how the end-users would utilise it. This was also assessed in the formative evaluation with end-users. The high degree of end-user involvement in the phases before and during the scenario based evaluation is important for the project.

Contribution to the TRADR scenarios and prototypes

This workpackage was responsible for defining the scenario and the other WPs gave their input to the scenario by incorporating the developed science and technology in the scenario. The Year 1 experimentation and evaluation will feed into the next year cycle to further refine and extend the scenario and use cases.

Furthermore, the scenario-based evaluation of the integrated system allows us to identify needs for further improvements with respect to systems development, combining the research objectives and the end-user perspective. The first year focused, among other things, on the identification of the (1) domain and organisational conditions and (2) stakeholders' values, needs and expectations (particularly, the fire brigade teams). In addition, specific support functions with expected effects (i.e., the claims) have been derived for inclusion in the scenarios (i.e., for design and evaluation).

1 Tasks, objectives, results

1.1 Planned work

The overall objective of WP7 in Year 1 is a scenario-based evaluation for multiple asynchronous sorties to assess a large-scale static disaster area (Task T7.1). This means that the team members operate under distributed control, with asynchronous communication. In order to support this, the following main aspects are needed:

- User needs analysis: the user needs analysis will uncover the needs when performing multiple asynchronous sorties to assess a large-scale static disaster area.
- Specification of the socio-technical design rationale: the requirements for the system have to be specified.
- Planning of scenario-based evaluation: planning of a scenario based evaluation, with a focus on enabling a fixed human-robot team to gradually build up situation awareness of a static disaster site over multiple, asynchronous sorties. Based on the user needs analysis for the Y1 scenario, the scenario needs to be defined with static events.
- Define methods and metrics: the methods and metrics have to be defined in order to assess performance during Y1's tasks.
- Assessment with end-users: assess performance of robots and end-users during multiple sorties in a large-scale static disaster area.

1.2 Actual work performed

In this section we describe the actual work that was performed, and how it feeds into the overall objective for WP7 in Year 1. We first describe a specification of the socio-technical design rationale in Section 1.2.1, through which the Y1 scenarios are defined. We then describe the various topics that were part of T-JEx, which are the end-user study (Section 1.2.2), reports created by the end-users regarding their T-JEx experiences (Section 1.2.3), Value Assessment Workshops (Section 1.2.4) and infield data collection with the Gaze Machine (Section 1.2.5).

1.2.1 Specification of the socio-technical design rationale

We apply the situated Cognitive Engineering (sCE) methodology [19], which is an iterative human-centred development process aiming at an incremental development of advanced technology. It consists of an iterative process of generation, evaluation, and refinement. Figure 2 shows the general structure of the sCE methodology, consisting of three components: 1) the *foundation* entails operational and human factors, and technological analysis to derive a sound and practical design rationale, 2) the *specification* provides maintenance of the requirements baseline, and 3) the *evaluation* through simulation or a prototype serves to validate and refine the requirements baseline. We describe these three components in more detail.

In the first component (foundation), foundational knowledge is described to identify actors, objectives, and contexts of the system and the (task) environment. Particularly, it describes the operational demands, the human factors (e.g. workload, situation awareness) and the technologies used (e.g. UGV and UAV, mobile tablet, operator workstation, command tactical display). This paragraph focuses on the first step to go from the operational demands (foundation) to a first design specification, i.e., the specification of the (design) scenario as part of the design rationale, and context for the evaluation.

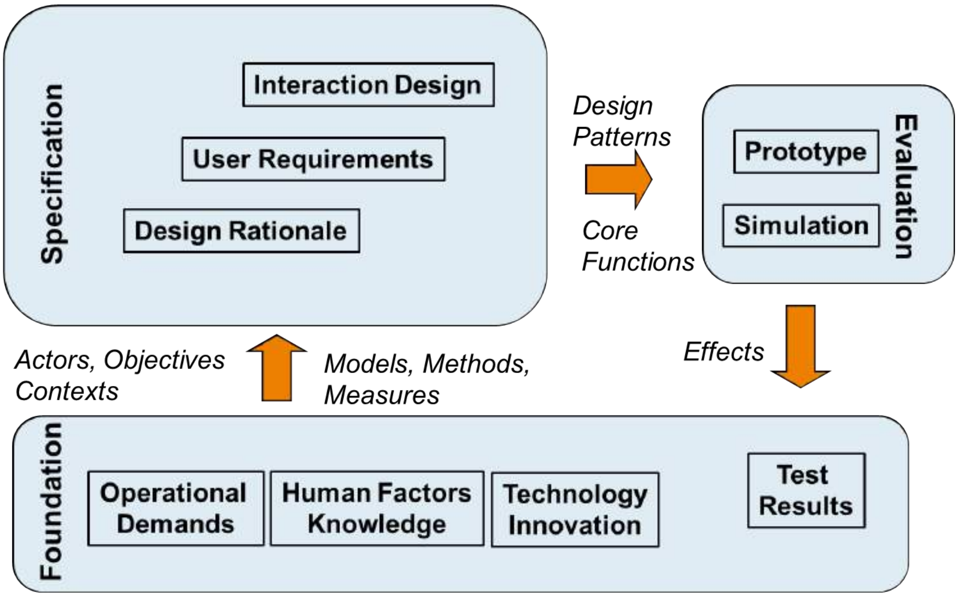

Figure 2: Schematic drawing of situated Cognitive Engineering method [19].

In Figure 3, the relations between aspects of the the specification phase are depicted in more detail. The list below explains the specification phase in more detail:

- Scenarios, which are part of the design rationale (see Figure 2) and give context to the use cases, support reasoning about situations of use, even before those situations are actually created. Scenarios can be elaborated as prototypes, through the use of storyboard (sequence of snapshots), video and rapid prototyping tools. From a general design scenario, more focuses and detailed use cases can be derived. Scenarios include:

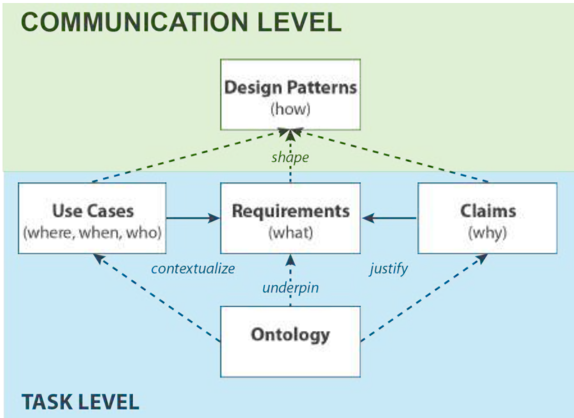

Figure 3: The sCE specification model.

a setting (starting state)

agents or actors, each actor has goals

a sequence of actions and events

- Use cases provide the (formal) contextualization (conditions, scope). Use-cases are short and structured prototypical examples of the envisioned man-machine interaction. From these use cases, we can derive requirements that describe what the machine should be able to do in order to make the use cases possible. Use cases illustrate the requirements.
- Requirements describe what the system should do
- Claims provide the justification (why). Claims specify the expected advantages and disadvantages of each requirement. They are used to justify the requirements by describing the effect on a certain measure that the requirement is expected to have. The claims come forth from the theoretic foundation (e.g., human factors knowledge).
- (Interaction) Design Patterns describe solutions to common usability or accessibility problems in a specific context.
- Ontologies provide a constrained vocabulary in which use cases, requirements and claims must be described.

The scenario and use cases have been defined by the whole TRADR consortium in close cooperation with end-users. In general they adhere to the following elements from the TRADR roadmap in Year 1: there are multiple asynchronous sorties in a large-scale static disaster with three human team members (team leader, UGV operator and UAV operator) and a UGV and UAV.

The use cases as defined by the TRADR consortium are the following:

- Initial overview of disaster area by UAV (Figure 4)
- Initial overview of areas of interest in disaster area and dispatch of UGV (Figure 5)
- Subsequent sortie and persistent information UAV/UGV
- Subsequent sortie and persistence information human team members (Figure 6)
- UAV building assessment
- UGV building assessment

Figure 4: Initial overview of disaster area by UAV (credits: Surveillance Drone by Luis Prado from The Noun Project).

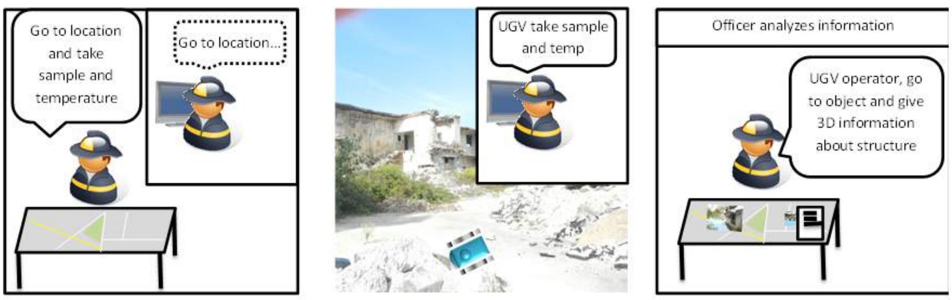

Figure 5: Initial overview of areas of interest in disaster area and dispatch of UGV.

In a workshop after the summer school (July 2014 in Prague), the TRADR team defined the requirements for the TRADR system according to four themes in the research:

- Human robot interaction

- Autonomy
- Persistence
- Team work

1.2.2 End-user study

To collect initial experience and data a scenario-based end-user study was conducted during T-JEx. The study was a qualitative study with three groups of end-users who were firemen from the VdF, Italy; to allow the participants to express themselves spontaneously, the study was conducted in Italian. Each group performed a task in the aftermath of an earthquake: look for dangerous substances and assess the state of the building. In Year 1, the study took place in an abandoned military hospital (Calambrone, Italy). For an overview of all test areas that are considered over the course of the TRADR project, see Appendix [6] (Annex Overview 2.7).

The story that provided context was that an earthquake occurred and the participants were asked to participate in the aftermath of the earthquake to look for dangerous substances and study the structural damage to the building. The victims had already been rescued by other firemen. The building they focused on was a military hospital, with suspected hazardous material present. The scenario that was prepared is shown in Figure 7. There were painted cracks in the wall, biohazard material and oxygen cans were lying around. The team consisted of three people, namely the team leader, the UGV operator, and the UAV operator. The UAV was piloted by a fourth person, which was not formally as a team-member, as his only task was to simulate an 'autonomous' UAV and had to translate the commands of the UAV operator to the UAV. The team had to find the cracks, biohazard material and oxygen cans and map them. On the outside, there were cracks that could be seen using UAV. Each day three participants each performed one of the following roles:

- Team Leader: command his team to assess the situation and make sure everything is safe.
- UGV operator: operating the ground robot and monitoring its video feeds.
- UAV operator: communicating with the UAV pilot to instruct where he has to fly. Monitoring UAV video feeds.

All team members used the TREX system as a common map (see the report on the "Overview of the development towards the TRADR Display System" in TRADR DR3.1 [5]). Here they could see where the team members and robots were and it gave them an overview of the situation. They could also add icons to the Tactical Display, such as photos taken from the UGV or UAV, remarks, dangerous substances etc. In addition, the exercise included a switch to a new team, in which each of the roles was taken on by a new participant; the old team had to update the new team of the current situation.

After executing of the mission, the fire fighters were debriefed during a discussion session on how they experienced working with the robot, which aspects they appreciated, and what could be improved.

In general, fire fighters see possibilities in the employment of robots, such as gaining information without being at risk themselves, and getting additional information from a different perspective. They also stated that the tactical display and information management could be improved, as for example only the team leader has the full overview, and that there are issues that make the employment of for example the UAV impractical, such as for example that the UGV still has little autonomy. The feedback will be used in the following year to improve the robot support in the end-user's urban search and rescue tasks.

1.2.3 End-user report on scenario based study

Representatives from the end-user partners were present at T-JEx and made observations about the exercise. Based on their experience they created a report, which details the T-JEx exercise from their respective viewpoints. The report from **IFR** can be found as Appendix [22] (Annex Overview 2.3) and the report from **VdF** can be found as Appendix [3] (Annex Overview 2.4). This formative evaluation identified the (1) specific domain and organisational conditions of the TRADR scenario and (2) stakeholders' values, needs and expectations (particularly, the fire brigade teams). In addition, specific support functions with expected effects (i.e., the claims) have been derived for inclusion in the scenarios (i.e., for design and evaluation).

1.2.4 Value assessment workshop

During T-JEx, three Value Assessment workshops were conducted with the end-users. The Value Assessment workshop is a technique that forms part of Value Sensitive Design, a user-centred design methodology [8]. In a Value Assessment workshop, end-users had to identify the stakeholder values of a particular technology.

Value Sensitive Design (VSD) is a user-centred design methodology that accounts for human values in a structured way throughout the design process. Key concepts in VSD are stakeholders, values and value tensions. In VSD, a distinction between direct and indirect stakeholders of a system is made. Direct stakeholders interact directly with the system or its output, and indirect stakeholders are impacted by the system without interacting

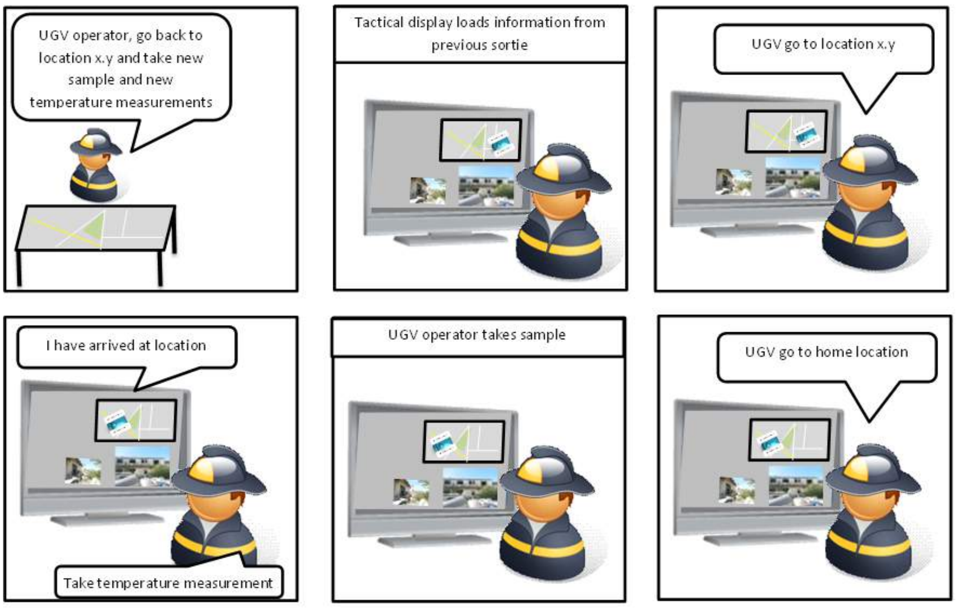

Figure 6: Subsequent sortie and persistence information human team members

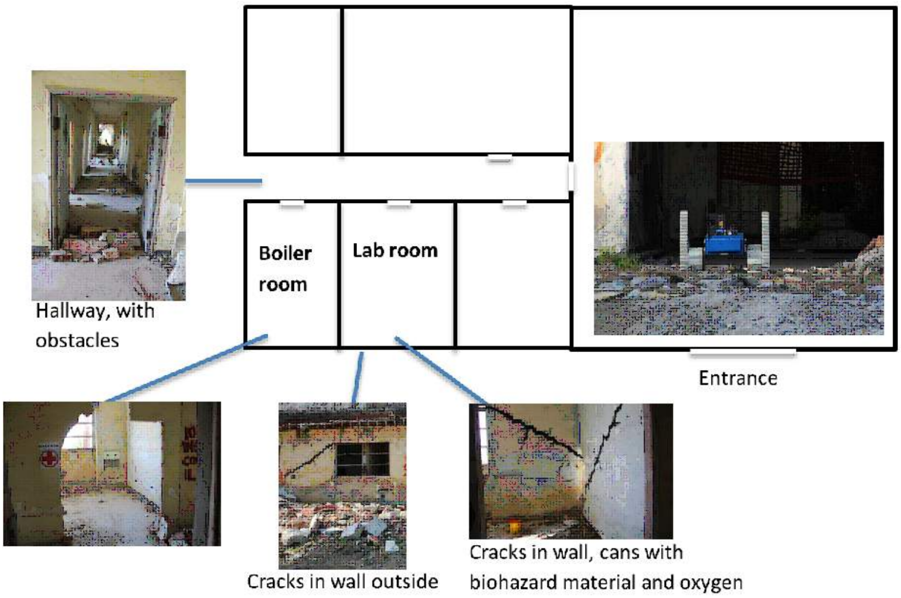

Figure 7: Schematic drawing of the scenario in military hospital.

with it directly. Stakeholders have values. A value refers to what a person or group of people considers important in life. Values that play a role in the design of technology are, for instance, autonomy, security, privacy, safety, trust, responsibility, sustainability, and fun. Value tensions occur when a particular design of a system supports one value, but hinders another. For example, supporting the value of security, e.g. by placing more surveillance cameras, may hinder privacy. VSD aims to make designers aware of these tensions during the design process so that they make informed design choices. The objective is to strive for improvements rather than perfection.

Figure 8: Working in groups during the Value Assessment Workshop.

In the workshops, the following direct and indirect stakeholders were identified by the end-users. Direct stakeholders are the fire-fighters; indirect stakeholders are the following: victims, paramedics, policemen, press, local authorities, curious observers, electricity companies, volunteers and insurance companies. The values identified for each stakeholder group can be found in the Appendix [10] (Annex Overview 2.5).

1.2.5 Infield data collection with the Gaze Machine

During T-JEx, we conducted a series of experiments using the Gaze Machine [25]. The Gaze Machine is a head mounted device which allows to estimate at each instant, where the person who wears it looks, in the 3D environment. The purpose of the experiments was to obtain a better understanding of the behaviour of the fire-fighters when they enter hazardous areas. During the experiment, one fire-fighter at a time walked through a part of the hospital wearing the GM. They were instructed to behave exactly as they would in an unknown environment after an accident, and comment on what they were doing and observing, as they would to colleagues in a command centre who do not have video feed from them.

Gaze data was collected made available to all the members of the consortium for further analysis and elaboration, e.g. analysis of verbal descriptions of unstructured and damaged environments in conjunction with gazing behaviour. Discussions with the participants have also highlighted various aspects which could help in introducing the GM as a tool capable to assist the fire-fighters in performing their demanding tasks. Moreover, the obtained reconstructions pave the way for the use of the GM for obtaining complete 3D maps of the environment from the data acquired during inspection or other compatible tasks. Finally, a detailed analysis of the acquired data will can the TRADR consortium to obtain a better understanding of the most crucial elements for structural and situation assessment. The full report of the experiment can be found as Appendix [20] (Annex Overview 2.6).

1.3 Relation to the state-of-the-art

Here we describe the relation to the state-of-the-art for each subsection of this deliverable in turn.

sCE and end-user evaluation. Development of interactive, human-centred automation should be built on theory and empirical research. To support the development processes systematically, recently, a situated Cognitive Engineering (sCE) method was constructed for building, maintaining and re-using design knowledge based on the following development principles [19]:

- Creating human-centred automation is a multi-disciplinary collaborative activity
- Functional modules are defined and tested incrementally in an iterative refinement process
- Design decisions are explicitly based on claims analyses, explicating the up-downside trade-offs
- Keeping and sharing the design rationale is key for progress and coherence in automation development

Corresponding to ISO 9241-210:2010 ("Human-centred design for interactive systems"), the situated Cognitive Engineering (sCE) methodology is an iterative human-centred development process of generation, evaluation, and refinement with intensive end-user involvement. It specifically aims at an incremental development of advanced technology [18], focusing on the development of joint (human-machine) cognitive systems, like the Cognitive Systems Engineering (CSE) approaches of Hollnagel and Woods [15]. The CSE-approach is combined with scenario-based design techniques that are common in usability engineering, emphasizing the specification of the design rationale for theory development [2]. The claims analysis part addresses the theoretical notions and provide a method for establishing a sound empirical validation. The scenario-analyses are part of a general cognitive work analysis with a focus on performance and resilience at the individual and team level [31],[14].

All the information generated by using the sCE method can be committed to the online sCE Tool[1]. The sCE method was successfully applied during the NIFTi and ALIZ-E projects.

Value assessment workshop. Value Sensitive Design as a framework for systematic accounting for human values in the creation of technology have been developed over the past 25 years; examples of human values that have been addressed are privacy, security, safety, responsibility, sustainability, autonomy and trust. For a thorough overview along with practical guidelines, see e.g.[8]. Despite the fundamental nature of human values in the creation of technology, methods like VSD have not been widely adopted. Borning and Muller [1] have suggested that this is the case because the VSD method has a tendency to *overclaim*, and propose some steps to remedy this.

Recently, some attempts have been made to integrate VSD with other design methods such as Requirements Engineering [11], and put VSD principles into practice in the design of support systems for train traffic control [12].

Gaze machine. The problem of gaze estimation has been given substantial attention in the literature. A survey regarding the problem is provided in [9]. The survey also illustrates that the problem is valuable in many areas of cognitive sciences and that still no definitive solution is available. Most 3D, model-based approaches [13, 28, 32, 35] rely on estimating the parameters of some model of the human eye. In [25] the authors propose a hybrid model considering both the geometric model of the eye and regression analysis, increasing the gaze estimation accuracy in three dimensions.

Regarding the importance of studying the points of regard (POR) in 3D scenes, it has been shown that the inclusion of depth perception changes human attention behaviour [34] and several attempts have been made to collect eye-tracking data while free-viewing stereo image pairs [16], artificially created 3D scenes [33], and even real 3D environments [27]. Limited options exist concerning systems for capturing PORs in 3D. The Gaze Ma-

[1] http://www.scetool.nl

chine (GM) [25] is a notable example, which uses high speed cameras to track the pupils of the subject's eyes and a stereo camera for the reconstruction of the scene. Due to the use of the stereo camera the device can work both indoors and outdoors. For indoor applications, RGB-D cameras have also been considered for the reconstruction of the scene [17]. Moreover, saliency estimation and prediction in 3D environments have recently gained an increased attention [24, 23, 21].

Regarding the study of eye movements in time critical scenarios and emergency situations, several neurological and psychological studies have been conducted. Numerous studies have been performed on the eye movements while driving (e.g. [29, 30]). Eye movements have also been studied in the context of military applications. For example pilot's eye movements inside the cockpit of military aircraft have been studied in [26] and [7]. Finally, a limited number of works have considered the eye movements of the operators in emergency response situations [4].

Using the Gaze Machine in TRADR contributes to the evaluation and further development of the GM, as the T-JEx sessions provide a realistic test environment in which the GM can be evaluated for use in USAR domains. In addition, it can help the TRADR consortium to obtain a better understanding of fire-fighters' behaviour during situation assessment.

2 Annexes

2.1 Smets, Nanja, Neerincx, Mark, Looije, Rosemarijn and Mioch, Tina (2014), "Specification of the socio-technical design rationale"

Bibliography Nanja Smets, Mark Neerincx, Rosemarijn Looije and Tina Mioch "Specification of the socio-technical design rationale" Unpublished technical report, TNO Soesterberg, 2014.

Abstract This document describes the use of the situated Cognitive Engineering Methodology and the specification of requirements and use cases for the TRADR system in detail. These were defined by the whole team in close collaboration with the end-users.

Relation to WP This report describes the specification of requirements and use cases important for evaluation with end-users in TRADR T7.1.

Availability Restricted. Not included in the public version of this deliverable.

2.2 Smets, Nanja (2014), "T-JEx end-user study TRADR Y1"

Bibliography Nanja Smets "T-JEx end-user study TRADR Y1" Unpublished technical report, TNO Soesterberg, 2014.

Abstract This document describes the TRADR end-user study. The study was conducted together with the TRADR Joint Exercise (T-JEx). The study was a qualitative study with three groups of end-users. The end-users had to perform a search for dangerous substances and look at the structural damage of a military hospital in the aftermath of an earthquake.

Relation to WP This report describes the study with end-users conducted during the TRADR Joint Exercise (T-JEx) T7.1.

Availability Restricted. Not included in the public version of this deliverable.

2.3 Pahlke, Norbert (2015), "IFR Pisa T-JEx observations"

Bibliography Norbert Pahlke "IFR Pisa T-JEx observations" Unpublished report, IFR, 2015.

Relation to WP This report describes T-JEx Y1 from the viewpoint of the IFR fire-fighters.

Availability Unrestricted. Included in the public version of this deliverable.

2.4 Corrao, Salvatore (2015), "Tradr Joint exercise - Report from Italian Fire Corps (CNVVF)"

Bibliography Salvatore Corrao "Tradr Joint exercise - Report from Italian Fire Corps (CNVVF)" Unpublished report, Italian Fire Corps (CNVVF), 2015.

Relation to WP This report describes T-JEx Y1 from the viewpoint of the VdF fire-fighters.

Availability Unrestricted. Included in the public version of this deliverable.

2.5 Harbers, Maaike (2014), "Value Assessment Workshops in the TRADR Project"

Bibliography Maaike Harbers. "Value Assessment Workshops in the TRADR Project." Unpublished technical report, Interactive Intelligence, TU Delft, 2014.

Abstract This document describes a series of three Value Assessment workshops that were conducted in the context of the TRADR project. In the first year of the TRADR project, a number of test sessions with end-users took place in Calambrone between 29 September and 1 October in 2014. During these days, three Value Assessment workshops were conducted. The Value Assessment workshop is a technique that forms part of Value Sensitive Design, a user-centred design methodology. In a Value Assessment workshop, end-users have to identify the stakeholder values of a particular technology.

This document contains an introduction into Value Sensitive Design, a description of the three Value Assessment workshops and an overview of the workshop results. The raw workshop results can be found at the appendices at the end of this document.

Relation to WP This report describes the stakeholder values important for defining the scenario in TRADR T7.1.

Availability Unrestricted. Included in the public version of this deliverable.

2.6 Ntouskos, Valsamis (2015), "Infield data collection with the Gaze Machine"

Bibliography Valsamis Ntouskos "Infield data collection with the Gaze Machine". Unpublished report, Sapienza University of Rome, 2015.

Relation to WP This report describes the experiments performed with the the Gaze Machine used by the fire-fighters during T-JEx. The purpose of the experiments was to obtain a better understanding of the behaviour of the fire-fighters when they enter hazardous areas.

Availability Unrestricted. Included in the public version of this deliverable.

2.7 de Greeff, Joachim, Corrao, Salvatore, Pahlke, Norbert and van den Broek Humphrey, Guido. (2015), "TRADR training and test areas."

Bibliography Joachim de Greeff, Salvatore Corrao, Norbert Pahlke and Guido van den Broek Humphrey "TRADR training and test areas." Unpublished overview, Interactive Intelligence, TU Delft, 2015.

Relation to WP This document provides an overview of the three test areas that are will be used for the yearly evaluations, over the course of the TRADR project.

Availability Unrestricted. Included in the public version of this deliverable.

References

[1] A. Borning and M. Muller. Next steps for value sensitive design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1125–1134, New York, NY, USA, 2012. ACM.

[2] J. Carroll and M. Rosson. Design rationale as theory. In *HCI Models, Theories and Frameworks*, pages 431–461. Elsevier Science, 2003.

[3] S. Corrao. TRADR Joint exercise. Technical report, Italian Fire Corps (CNVVF), Italy, 2015. Unpublished.

[4] D. Crundall, P. Chapman, N. Phelps, and G. Underwood. Eye movements and hazard perception in police pursuit and emergency response driving. *Journal of Experimental Psychology: Applied*, 9(3):163, 2003.

[5] J. de Greeff. Overview of the development towards the TRADR Display System. Technical report, Delft University of Technology, 2015. Unpublished.

[6] J. de Greeff, S. Corrao, N. Pahlke, and G. van den Broek Humphrey. TRADR training and test areas. Technical report, Interactive Intelligence, TU Delft, 2015. Unpublished.

[7] P. M. Fitts, R. E. Jones, and J. L. Milton. Eye movements of aircraft pilots during instrument-landing approaches. *Ergonomics: Psychological mechanisms and models in ergonomics*, 3:56, 2005.

[8] B. Friedman, P. H. Kahn Jr, A. Borning, and A. Huldtgren. Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory*, pages 55–95. Springer, 2013.

[9] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(3):478–500, 2010.

[10] M. Harbers. Value assessment workshops in the tradr project. Technical report, Interactive Intelligence, TU Delft, 2014. Unpublished.

[11] M. Harbers, C. Detweiler, and M. A. Neerincx. Embedding stakeholder values in the requirements engineering proces. In *Proceedings International Working Conference on Requirements Engineering (REFSQ)*, 2015.

[12] M. Harbers and M. A. Neerincx. Value sensitive design of automated workload distribution support for traffic control teams. In *Proceedings*

of the International Conference on Human Computer Interaction, pages 12–21. Springer, 2014.

[13] C. Hennessey, B. Noureddin, and P. Lawrence. A single camera eye-gaze tracking system with free head motion. In *Proc. symp. on Eye tracking research & applications*, pages 87–94, 2006.

[14] E. Hollnagel, J. Paris, D. Woods, and J. Wreathall. *Resilience Engineering in Practice.* Ashgate, 2011.

[15] E. Hollnagel and D. Woods. Cognitive systems engineering: new wine in new bottles. 18(6):583–600, 1983.

[16] C. Lang, T. V. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan. Depth matters: Influence of depth cues on visual saliency. In *Computer Vision–ECCV 2012*, pages 101–115. Springer, 2012.

[17] C. McMurrough, C. Conly, V. Athitsos, and F. Makedon. 3d point of gaze estimation using head-mounted rgb-d cameras. In *Proc. Int? ACM SIGACCESS conference on Computers and accessibility*, pages 283–284, 2012.

[18] M. A. Neerincx. Situated cognitive engineering for crew support in space. *Personal Ubiquitous Comput.*, 15(5):445–456, June 2011.

[19] M. A. Neerincx and J. Lindenberg. Situated cognitive engineering for complex task environments. In *Naturalistic Decision Making & Macrocognition*, pages 373–390. Ashgate Publishing Limited, 2008.

[20] V. Ntouskos. Infield data collection with the gaze machine. Technical report, Sapienza University of Rome, 2015. Unpublished.

[21] V. Ntouskos, F. Pirri, M. Pizzoli, A. Sinha, and B. Cafaro. Saliency prediction in the coherence theory of attention. *Biologically Inspired Cognitive Architectures*, 5:10–28, 2013.

[22] N. Pahlke. IFR Pisa TJEx observations. Technical report, IFR, 2015. Unpublished.

[23] L. Paletta, K. Santner, G. Fritz, H. Mayer, and J. Schrammel. 3d attention: measurement of visual saliency using eye tracking glasses. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 199–204. ACM, 2013.

[24] H. S. Park, E. Jain, and Y. Sheikh. 3d social saliency from head-mounted cameras. In *Advances in Neural Information Processing Systems*, pages 431–439, 2012.

[25] F. Pirri, M. Pizzoli, and A. Rudi. A general method for the point of regard estimation in 3d space. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 921–928, 2011.

[26] M. G. Sanders, R. R. Simmons, and M. A. Hofmann. Visual workload of the copilot/navigator during terrain flight. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 21(3):369–383, 1979.

[27] K. Santner, G. Fritz, L. Paletta, and H. Mayer. Visual recovery of saliency maps from human attention in 3d environments. In *Int'l Conf. Robotics and Automation*, pages 4297–4303, 2013.

[28] S. W. Shih and J. Liu. A novel approach to 3-d gaze tracking using stereo cameras. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(1):234–245, 2004.

[29] M. Sodhi, B. Reimer, and I. Llamazares. Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods*, 34(4):529–538, 2002.

[30] G. Underwood. Visual attention and the transition from novice to advanced driver. *Ergonomics*, 50(8):1235–1249, 2007.

[31] K. J. Vicente. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Lawrence Erlbaum Associates, New Jersey, April 1999.

[32] A. Villanueva and R. Cabeza. A novel gaze estimation system with one calibration point. *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, 38(4):1123–1138, 2008.

[33] D. Vishwanath and E. Kowler. Saccadic localization in the presence of cues to three-dimensional shape. *Journal of Vision*, 4(6):4, 2004.

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

[35] Z. Zhu and Q. Ji. Novel eye gaze tracking techniques under natural head movement. *IEEE Trans. Biomedical Engineering*, 54(12):2246–2260, 2007.

IFR Pisa TJEx observations

Some critical aspects to the Pisa test

The test in Pisa emphasised that the backbone of the system is a well working communication infrastructure with a sufficient bandwidth and range. Otherwise latencies will slow down the response time or a broken connection will provoke the loss of a robot, namely an UGV. A communication tool like the TREX/TDS is very helpful for the end-users in sense of situation awareness followed by situation management. The better overview along the chain of command achieved by live photos/videos and positioning information rises the effectiveness in team management. It will be strongly influenced by the design of the tool. The experience from Pisa offered problems concerning clearness of presentation. A listing with different points was created.

Tradr Joint exercise, period 22 September to 3 October 2014 – Report from Italian Fire Corps (CNVVF)

Tradr Joint exercise operational phases on field took place in a former US military hospital (US AIR FORCE IN ITALY) decommissioned on 70's , located in Calambrone, Municipality of Pisa.

The large building was partly appropriately demolished to obtain a site as close as possible to the reality of buildings affected by an earthquake or explosion. There, conditions of instability and voids with secondary accesses have been created to further place actors simulating victims without being seen by rescuers and dogs units. This to avoid follow their tracks when access to rubble.

The site is normally used for exercise purposes of Medium USAR module, based at Pisa Provincial Fire station.

To let all participants know about an USAR team intervention during the Joint Exercise, a real search and rescue operation has been simulated and performed by Pisa M (Medium)[1] - USAR team, highlighting the strategies of intervention in case of collapses, such as:

- Information management;

- Areas of work zoning;

- Building static assessment;

- Presence of plants and hazardous substances assessment (HAZ MAT);

- Search;

- Rescue.

During the simulation, main aspects related to K9 time management and work methodology and listening electronic equipment have been presented, highlighted and deployed. Marking and communication system in USAR environment were also shown as reported in INSARAG international guidelines. Operational support documents were also presented and displayed.

During the joint exercise, special skills, capabilities and capacities from the Italian Fire Corps were present on the field, such as k9, CBRN, USAR and HELICOPTER units. Personnel from these sectors had the chance to observe technologies displayed like the gaze machine, UAVs, UGVs, bionic arm in order to get an idea about their future applications in rescue activities and provide suggestions about abilities to enhance in these devices.

USAR applications

UAVs and UGVs could give a good support for static and haz-mat assessment activities to define site access by the USAR team in the early stages of a scenario evaluation. This would help operators to avoid exposures to possible hazards due to further collapses in buildings already damaged. Robots would offer the chance to investigate parts not easily accessible for height and/or other unstable, risky conditions. Robots would avoid using means much more challenging and sometimes impossible to deploy.

[1] http://www.insarag.org/

Use of drones (UAVs) is also important to have an integral vision (better overview) from above of the working areas and surrounding conditions. This would avoid the classic problem of actual helicopters intervention that can create great air movement and noise provoking further collapses in buildings that have precarious stability conditions. Use of drones would also limit operational costs during rescue.

Another use could be to follow and track dogs during a search when the conditions do not allow handlers to follow their dogs during search operations.

UGVs, despite having a considerable volume, could offer great opportunity of use especially to assess potentially hazardous areas and not immediately accessible, for instance industrial facilities damaged .

They are useful to define positions where hazardous materials are present, inspect collapsed sites, verify cracks without accessing into them directly. The possibility to apply chemical and radiological detectors make also robots undoubtedly a valuable aid in high residual risk conditions.

Value Assessment Workshops in the TRADR Project

Maaike Harbers, Interactive Intelligence, TU Delft
October 2014

1. Introduction

This document describes a series of three Value Assessment workshops that were conducted in the context of the TRADR project. TRADR is an integrated research project funded by the EU FP7 Programme on long-term human-robot teaming for robot-assisted disaster response. The aim of the project is to develop technology for human-robot teams to assist in disaster response efforts, over multiple missions. A user-centric design methodology is used for the development of this technology.

In the first year of the TRADR project, a number of test sessions with end-users took place in Pisa between 29 September and 1 October in 2014. During these days, three Value Assessment workshops were conducted. The Value Assessment workshop is a technique that forms part of Value Sensitive Design, a user-centered design methodology. In a Value Assessment workshop, end-users have to identify the stakeholder values of a particular technology.

This document contains an introduction into Value Sensitive Design, a description of the three Value Assessment workshops and an overview of the workshop results. The raw workshop results can be found at the appendices at the end of this document.

2. Value Sensitive Design

Value Sensitive Design[1] (VSD) is a user-centered design methodology that accounts for human values in a structured way throughout the design process. Key concepts in VSD are stakeholders, values and value tensions.

In VSD, a distinction between direct and indirect *stakeholders* of a system is made. Direct stakeholders interact directly with the system or its output, and indirect stakeholders are impacted by the system without interacting with it directly. Stakeholders have values. A *value* refers to what a person or group of people considers important in life. Values that play a role in the design of technology are, for instance, autonomy, security, privacy, safety, trust, responsibility, sustainability, and fun. *Value tensions* occur when a particular design of a system supports one value, but hinders another. For example, supporting the value of security, e.g. by placing more surveillance cameras, may hinder privacy. VSD aims to make designers aware of these tensions during the design process so that they make informed design choices. The objective is to strive for improvements rather than perfection.

The VSD methodology contains three parts: conceptual, empirical and technical investigations. Conceptual investigations involve the analysis of direct and indirect stakeholders, their values, and how the envisioned technology affects their values. Empirical investigations involve the elicitation of stakeholders' views and values, and evaluations of prototypes. Technical investigations involve the assessment of existing technologies and solutions, and the development of prototypes.

[1] Friedman, B., Kahn Jr, P. H., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. In Early engagement and new technologies: Opening up the laboratory (pp. 55-95). Springer Netherlands.

Value Sensitive Design knows a number of methods and techniques that support the conceptual and empirical investigations. The Value Assessment workshops conducted at the end-user meeting in Pisa were strongly inspired by the Value Story method[2]. The next section describes the specific steps and activities that were performed in each workshop.

3. Workshops

The basic steps in all three workshops are: a) identify stakeholders, b) identify values, and c) examine the relation between stakeholder values and technology. In order to perform step b and c, one always has to identify stakeholders first (step a). Subsequently, there are two possible ways to proceed. The first is to immediately start identifying values for each stakeholder groups (step b), and then explore how technology affects the identified values (step c). A second way is to start with a particular (envisioned) system and, for each stakeholder group, identify how the technology can harm and benefit different stakeholders (step c). Then, the harms and benefits can be mapped onto stakeholder values (step b).

In workshop 1, the basic steps were performed in the order a-b-c, and in workshop 2 and 3, they were performed in the order a-c-b. In workshop 1, a number of steps were added to the basic steps, and in workshop 2 and 3, one step was omitted for reasons explained below.

Workshop 1

Date:	Monday 29 September
Time:	8:30-12:30 (4 hours)
Participants:	12 firefighters at different levels in the organization (both officers and field workers)
Timing:	the participants had not observed or participated in a test session yet

Workshop program:

1. Create a mind map of "Disaster Response"
2. Attend a presentation about values and technology
3. Identify direct and indirect stakeholders
4. Identify values of stakeholders, for each value give at least one concrete situation
5. Identify stakeholder needs for the concrete situations
6. Prioritize the importance of stakeholder needs (low, medium or high)
7. Discuss conflicts between values and stakeholder needs

Step 3, 4, and 5 (identify stakeholders, identify values, examine relation between values and technology, respectively) form the basic steps of this workshop. The first step is a warming up and was added to engage the workshop participants. Step 2 serves to motivate the Value Assessment workshop by explain why values are important in the design of technology. Furthermore, the presentation aims to prepare the workshop participants for step 4. Step 6 and 7 were added to obtain additional information from the participants regarding priorities of different values and possible conflicts between different values.

Workshop 2

Date:	Tuesday 30 September
Time:	13:30-15:30 (2 hours)
Participants:	8 firefighters, mostly field workers
Timing:	all of the participants attended workshop 1 the day before, and after that, they

[2] Detweiler, C., & Harbers, M. Value Stories: Putting Human Values into Requirements Engineering.

observed and/or participated in two test sessions

Workshop program:

1. Identify harms and benefits of the technology for each stakeholder group
2. Map harms and benefits onto corresponding values

The participants of workshop 2 already followed workshop 1 on the day before. Therefore, we skipped the first basic step (step a), identifying stakeholders, and used the list of stakeholder groups the participants identified in the previous workshop.

Workshop 3

Date: Wednesday 1 October
Time: 10:00-11:30 (1.5 hours)
Participants: 3 firefighters, field workers
Timing: the participants had not observed or participated in a test session yet

Workshop program:

1. Identify direct and indirect stakeholders
2. Identify harms and benefits of the technology for each stakeholder group

In this workshop, the step of mapping harms and benefits onto values was omitted due to time constraints.

4. Results

The results presented here are the merged results of all three workshops. In the appendix at the end of this document, the raw results of the separate workshops are provided.

In the workshops, the following direct and indirect stakeholders were identified.

Direct stakeholders

- Firefighters

Indirect stakeholders

- Victims
- Paramedics
- Policemen
- Press
- Local authorities
- Curious observers
- Electricity company
- Volunteers
- Insurance companies

In this section, the identified values for each stakeholder group are provided, together with how the envisioned robot technology could possibly support and/or hinder these values.

Firefighters

Value	Relation between value and technology	
	Support	**Hinder**
Personal safety	Robots make it possible for firefighters to stay away from dangerous situations.	
Safety of others	Robots provide extra technical support that make it possible to search and rescue victims (humans and animals) faster.	
Access to information	Operators can see more when using robot technology, and robots make it possible to explore places that are normally inaccessible. This leads to a better situation awareness.	It can happen that the operator loses radio contact with the robot, and the operation is temporarily suspended.
Well-being	Firefighters maintain more emotional distance when they see victims via a robot.	
Effectiveness	Robot technology can make a firefighter more effective if it functions well, e.g. by enabling better situation awareness.	Robot technology has weaknesses that can decrease effectiveness, e.g. robots only functions in good weather circumstances, the radio signal isn't always present (e.g. under rubber), technology for making high resolution photos is very slow, robot carries less special equipment.
Ease of use	Firefighters can work more effectively with robots that are easy to use.	
Authority	Robots can be used to have the best situation awareness compared to other organizations.	If there are mixed teams (fire brigade, robots, other organizations) or if there are communication barriers with robots, this can lead to decreased authority for leading firefighters.

Victims

Value	Relation between value and technology	
	Support	**Hinder**
Personal safety	Robots can find and rescue victims.	Robots can be dangerous for victims, e.g. if they have inflammable batteries, or when a drone flies into a human.
Health	Robots can provide first aid to victims, e.g. provide medicine or water with mechanical arms, and they can obtain information with sensors about their medical situation.	

Well-being	The robot can transmit an ‘indirect’ human presence, and they can establish instant communication with a human search and rescuer.	The effect of seeing a robot as a victim can be shocking. Robots cannot provide the same level of reassurance as humans.
Access to information	Robots can provide information to victims when they are lost, e.g. a map with information, camera pictures, radio.	
Contact	Robots can establish contact between victim and others through radio, camera, a monitor, a way to input information.	

Paramedics

Value	**Relation between value and technology**	
	Support	**Hinder**
Personal safety	Robots can take risks away from paramedics because they can stay away from dangerous areas, e.g. when there are chemical substances, radioactivity, extreme temperatures.	
Access to information	Paramedics can obtain more information through robots, e.g. when they have to perform triage or estimate the number of victims and their situation.	
Contact	Robots allow paramedics to talk to victims they would not be able to talk to otherwise.	It is more difficult to establish a relation with a victim through a robot or drone.
Health	Robots can provide information about the physical state of victims, e.g. circulation, breath, heart rate.	Robots cannot provide health information about a victim of the same quality as a human could.
Well-being		Paramedics can do less for well-being of victims because they can only give reassurance by radio and it is more difficult to assess of the victims vital functions due to lack of physical contact.

Policemen

Value	Relation between value and technology	
	Support	**Hinder**
Personal safety	It is safer for a policeman to control a robot to go to a dangerous area himself, e.g. when public order is disrupted, disasters.	
Security	Robots provide extra technical support to create security, e.g. being monitored prevents criminality.	Robots may make it more difficult for policemen to stay alert during task performance (vigilance tasks).
Neutrality	When robots are safe and secure they can help policemen in being neutral, e.g. during strikes.	
Effectiveness	Robots provide extra technical support to create security, e.g. increased situational awareness.	Robots may not work under all conditions, this can make police less effective.
Courage	When robots are safe and secure they can help policemen in showing leadership, creating a good work environment with colleagues and population, and creating a good exchange of information with other stakeholders.	
Security	When robots are safe and secure they can help policemen in creating security, e.g. when patrolling an area or when cooperating with the public.	
Trust	When robots are accepted by users and publics, they may increase trust in police.	
Access to information	Robots can provide more information to policemen, e.g. during investigations, special operations, intervening people, etc.	

Press

Value	Relation between value and technology	
	Support	**Hinder**
Impartiality	Robots collect neutral information that press can use.	Robots that are not secure threat impartiality, e.g. hackers, spam, viruses.
Transparency	Robots can give more transparency regarding the current situation.	Robots that are not secure threat transparency, e.g. hackers, spam, viruses.
Access to information	Robots allow press to transmit more information and faster to public and other stakeholders, e.g. photos or videos of situation.	Extra information may also alarm population unnecessarily.

Local authorities

Value	Relation between value and technology	
	Support	**Hinder**
Access to information	Robots can give local authorities extra access to information, e.g. through camera, GPS, and mapping and presentation software.	
Sharing information with others	Robots allow local authorities to provide more information about the situation to press, citizens and family of victims.	
Safety	Robots provide extra support for public safety, e.g. presence of robots with cameras during a disaster can decrease criminality.	
Healthy finances		Robots may yield high costs.

Curious observer

Value	Relation between value and technology	
	Support	**Hinder**
Curiosity	It can be interesting to observe a robot in function.	Curious observers may hinder operations with robots.
Safety	Robots (UAVs) can be used to monitor curious observers of a disaster and detect dangerous situations.	

Electricity company

Value	Relation between value and technology	
	Support	**Hinder**
Access to information	Robots allow electricity company to monitor the area, investigate a particular damage	
Safety	Robots can make it safer for employees from the electricity company to examine infrastructure	

5. Value tensions

The aim of the TRADR project is to develop new technology for long-term human-robot teams for disaster response. The previous section showed an overview of the stakeholders of that technology, their values, and how the technology can support and hinder those values. If supporting one such value hinders another value, there is a value tension. Value tensions can involve values of different stakeholders groups, different values of one stakeholder group, or one value of one stakeholder group. In this section, a number of possible value tensions regarding the design and deployment of UGVs and UAVs will be discussed. Some of the value tensions were identified by the end-users in workshop 1, and others are based on an analysis of the workshop results.

Hindering vs supporting safety

Robots can both support and hinder the safety of the people that encounter them, such as victims and rescue workers. On the one hand, robots can improve the search and rescue operation. But on the other hand, they can be dangerous, for instance, when they fail to identify a human being and collide (flying or driving) with the human. Also, equipig robots with weapons and amunition may support the safety of search and rescue workers or policemen, but may hinder the safety of victims or other people encountered by the robot.

Safety vs well-being

The deployment of robots can support safety of victims by making the search and rescue operation faster and more effective, but it can hinder the victims' well-being. For example, it may be a shocking experience to be trapped, wounded and lost, and suddenly be confronted with a robot, in particular, if there are no humans around. There may also be victims that do not like robots or that do not want to be saved by a robot.

Effectiveness of firefighter vs police

Search and rescue robots can be deployed for a lot of different activities. When there is a limited amount of robots, choices have to be made regarding their deployment. In such a situation, for instance, if a robot is deployed for activities of the fire brigade, effectiveness of firefighters is support the effectiveness of policemen is hindered, and vice versa if the robot is deployed for activities of the police. This tension may also occur within one stakeholder group, e.g. firefighter, when the group is divided in subteams, and there are not sufficient robots for all subteams.

Transparency vs privacy

Robots make it possible to collect more information of a disaster through their cameras and other sensors. Transmitting this information to the press supports transparency, but it may hinder privacy of victims. On the one hand, press can better inform the public about the situation at hand. But on the other hand, it may happen that privacy sensitive information about victims is spread this way. For instance, family members learn about a victim's situation through media rather than through personal conversation.

Safety and effectiveness vs healthy finances

Deployment of robots can increase the safety and effectiveness of rescue workers during a disaster response situations. However, the purchase of robots may be expensive and hinder the local authorities value of healthy finances.

Transparency and access to information vs well-being

Robots make it possible to collect more information of a disaster through their cameras and other sensors. Spreading this information can support transparency and access to information for the public and other stakeholders. But at the same time, it may hinder well-being by scaring people and create unnecessary panic.

Appendix

Mind map of "disaster response"

Working in groups

Working in groups

Results of workshop 1

Workshop 1

Stakeholder	Value	Situation	Need	Priority		
				L	M	H
Firefighter	Personal safety	Rescue flight with onboard squad of saf2b unit and n.b.c.r. unit, u.a.v. pilot, chemical plant fire with high towers	In chemical environment: robot with the right instrument can air sampling, make video/infrared, surveillance			x
	communication	During a rescue operation we lost the radio communication and temporary suspension of the operation	The robot will be useful as video/communication (radio), mobile device, and to see the rescue squad in order to maintain a correct situation awareness, and give us radio informations			x
	leadership	Assessment of the scene, task assignment, manage and control the work is made good with safety, car crash with people to save	A wire guided U.a.v. flies aove the scene and gives at the team leader more infomration about the scenario from above			X
	teamwork	Earthquake: during a main disaster with combined teams of other organization, there can be lack of leadership authority and commnication barriers.	A robot can be used to have a best situation awareness also for the other organizations on the scene, e.g. stay out of the red zone, building collapse imminent, etc.			x
Victims	Personal safety	When in a dangerous situation	Radio, camera, sensors that can help to increase situational awareness			x
	Contact	When lost, trapped, wounded	Radio, camera, monitor with movement sensor, writing system, infomration input			X
	Health/well-being	When in need for medicine, wounded	Mechanical arm to give medicine or water, sensors to give medical assistance		X	
	Guidance	When lost	Radio, camera, map with information			X
	Reassurance	When in panic, stressed	Radio, camera/visual contact		X	
Paramedics	Personal safety	When having to communicate with other incident responders	Camera, sensors for: Ex/Ox, LEL, chemical substances, radioactiviey, temperature			x
	Access to information	When in triage situation, needing info about number of victims and their situation	Camera (thermal), GPS, mapping + location, photo + GIS information			X
	Contact	When needing to talk with victims	Radio (voice controllerd), camera for visual contact			X
	Health	When needing to know biological situation of a victim, e.g. continuous, circulation, breath	Thermal camera, visual camera, night vision, sensors for: breath/fumes, detect life/death, circulation			x
Policemen	Personal safety	When public order is disrupted, e.g. during strikes, mass gathering events such as concerts, terrorist attacks	Robots have to be ruggedized, resistant, high durability for long term operations, timing, not harmful to operators			x
	Neutrality/ impartiality	During strikes, to respect vans, any time situation, presence of policemen is a determant of legality	Robots have to be safe and secrue (access code, no viruses, no spam, wifi protected, safe cloud system)		x	x

	Courage	- When showing leadership - when creating a good work environment with colleagues and population - when creating a good exchange of information with other stakeholders - when conveying trust and security to people	Robots have to be safe and secrue (access code, no viruses, no spam, wifi protected, safe cloud system)			
	Security	- when patroulling in an area - when cooperating with public	Robot with camera to inspect dangerous areas, gun, special glasses, night camera/vision, special sensors working autonomously, UAV patrolling and rescuing in a hazardous area		x	
	Trust	- when being trustworthy for collegues and public - public instituations have to trust policemen	Efficiency in robots/devices, and robots must be accpted by users + public			X
	Stratigic/ intelligence skills	During investigation, special operations, intervening people, capacity to build and cathc samples, infomarion and find something useful for investigation	Skills to collect data and analysis, data storage, capacity to record data and make comparisons, durability in long term operation, data assessment, dynamic mapping, network, high data transmission from user to robot and back		x	X
Press	Impartiality	Information I give has to be impartial as I am an instrument/tool for disaster response	Robots have to be safe/secure, not allowing for hackers or spam, security access codes to log in and out, 1 client authorized			x
	Transparancy	Information must be transparant, without influence of any kind	Robots have to be safe and secrue (access code, no viruses, no spam, wifi protected, safe cloud system)			x
	ICT skills/ ability	Use high technology to transmit data and info to public but also stakeholders	Robots must be compatible with updated technology and with my network + devices, integration, user friendly, ready to be deployable, ready to be used, not to interfere with other devices, has to come back home when it loses connection, easy to be washed and decontaminated without breaking it, to be used in any weather conditions, easy to be used by my smartphone, cost effective, firefighter proof			x
Local authorities	Access to information	- When they have to make a plan and inform stakeholders - When they have to make decisions	Camera, GPS, mapping and presentation software		X	
	Sharing information	Inform families of victim, press and citizens	Camera, GPS, mapping and presentation software		X	

Workshop 2								
	Firefighter	**Victims**	**Paramedics**	**Policemen**	**Press**	**Local authorities**	**Curious observer**	**Electricity company**
Benefits	Operator can stay outside, which is safer	Transmit an 'indirect' human presence	Identify a safety and operative area for paramedics operations	Safety for operator	Take a photo or video for a true depiction of situation	Extra technical support for public safety	Monitoring by video or photo the situation	Area monitoring
	Operator can see more		Monitoring victims/ people presence and observe their health	Extra technical support for security	Smart exchange information	More information for decision	Monitoring the c.o. presence	View of a particular damage
	Possibility to explore places you normally cannot go			Being monitored prevents criminality	Speedy propagation of the information to local people	Prevents criminality		Test an electrical device or structure
	More emotional distance when seeing a victim							Safety for operator
	Extra technical support for rescue							
Harms	Slowness of equipment	The drone could hurt human	The drone isn't able to test the human health	Only works with good weather	The use of the drone for smart exchange information can alarm the population	Needs more autonomy	Could catch curious observer attention	The robot actually is limited for some tasks
	Only works with the good weather	To give a wrong evaluation about the presence or not of human/ victim in the scenario	The drone cannot establish a relationship with human	Difficult to remain alert for operator	Catch attention of curious observer	Expensive		
	High resolution photo only in post-production (slow)	It cannot transmit a level of confidence or other benefit as a human rescuer						
	Dependent on radio signal (does not work under rubber)							
	Robot carries less special equipment							

Workshop 3			
	Firefighter	**Victim**	**Paramedics**
Benefits	Good/better situational awareness	To have a robot to find and save you	Take risk away from paramedics (they can stay outside)
	No risks for operators (CBRN – Chemical Bio Nuclear Radiological)	Instant communication with helper	Visualize victim's conditions
	To rescue animals	Together with operator bring first aid	Give vital information about victim
	To collect data remotely		
	Aerial survey before working in hazard areas -> survey assessment		
Harms	Robots can break so my mission ends with these machines	Effect of seeing a robot	Lack of physical contact with victim
	Less authority on a long mission(battery ends after a short/medium time)	Inflammable battery	Bas assessment of vital functions
			Give reassurance only by radio

Infield data collection with the Gaze Machine

Valsamis Ntouskos
Sapienza University of Rome

October 2014

Abstract

This report describes the experiments performed with the use of the Gaze Machine, during the last three days of the first TRADR joint exercise. The experiments regarded the behaviour of the firefighters in the task of structural assessment and overall inspection of a damaged, and partially collapsed structure. The experiments lead to interesting discussion with the end users about the use of the Gaze Machine in their field and produced useful data which can be used by the members of the consortium for addressing relevant problems.

1 Introduction

During the first TRADR joint exercise (TJEx2014), we conducted a series of experiments using the Gaze Machine (GM), shown in Fig. 1. The purpose of the experiments was to obtain a better understanding of the behaviour of the firefighters when they enter hazardous areas, like in this particular case, a partially collapsed building.

In the context of the use case considered for TJEx 2014, the experiments' site simulated a hospital facility after an explosion which had as a result a partial collapse of the building. Moreover, the presence of biological and chemical hazards was suspected. In this setting, firefighters volunteered to enter the site from an opening, wearing the GM, and inspect the building. In particular, they were asked to simulate a real-time communication with the command post, providing information they considered crucial for subsequent sorties, as well as an overall assessment of the situation inside the building. This assessment included the presence of dangerous materials, other types of hazards as well as an assessment of the structural elements of the building. The details of this particular scenario have been decided after fruitful conversations between the firefighters and TRADR members.

2 Gaze Machine

The Gaze Machine is a head mounted device which allows to estimate at each instant, where the person who wears it looks, in the 3D environment. It is

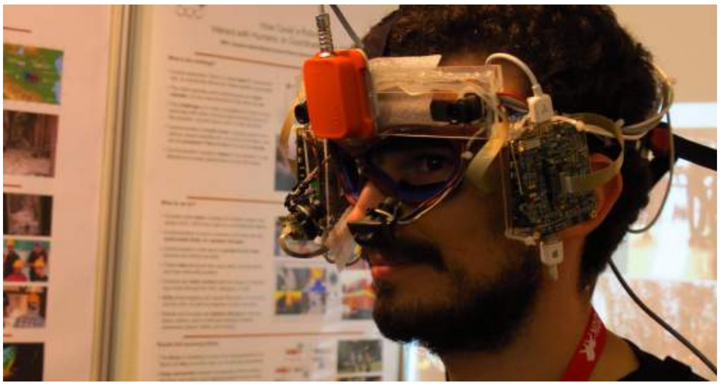

Figure 1: The Gaze Machine

composed by three cameras. One high speed infrared camera tracks the pupil of the subject in real-time via specialized software. The other two cameras capture the scene in front of the subject that uses the GM. The images from these cameras form a stereo pair and are used to obtain a 3D reconstruction of the scene. By means of a calibration procedure, the point of regard (POR), i.e. the point in space where the person wearing the GM looks at, is recovered both in the images of the scene and in the 3D reconstruction. The Gaze Machine can be used in many applications. Examples include research and scientific studies (e.g. computational attention, psychology, psychophysiology), medical diagnosis of various conditions and syndromes as well as Augmented Reality and Robotics. For a detailed description of the gaze estimation model used by the GM see [1].

3 Experiment description

The experiments were divided in three phases.

During the first phase, a detailed description of the Gaze Machine and the data we are interesting to collect, was provided. Based on this, specific guidelines about the correct use and protection of the GM were given to each participant. After a brief discussion, where possible issues were addressed, the participant wore the GM and was assisted in calibrating it. This concluded the first phase.

During the second phase, the participant was accompanied to the opening from where she/he had to enter the structure. There the following description of the task that they had to perform was provided: “You are asked to enter the hospital for inspection of the overall conditions. Describe what you see as you navigate through the site, simulating a real scenario where you have to report to a colleague what you see. Focus on structural assessment of the building and on any hazards you discover. The path you should follow will be described by the coordinator of the Firefighter’s team.”

After the details about experiment execution were provided, data recording of the GM was activated and the firefighter entered the structure. At this point no further contact with the participant was possible until he concluded the inspection of the site. With the exit of the participant from the site, recording was halted and before removing the GM a second calibration was performed in order to be able to assess the validity of the data. The duration of each experiment is given in Table 1. The duration of the recorded data for each experiment was around 800 [sec], with the exception of FF 3 where the recording was interrupted shortly after the experiment's start, most probably due to a hardware communication issue caused by some impact.

At the third and final phase, an interview with the participant was conducted in order to collect information and opinions about the progress of the experiment, the use of the GM and its possible applications in their field. The answers of the participants are provided in Appendix A.

Figure 2: Firefighter wearing the GM enters the site through an opening

4 Results

The data collected at the site were subsequently processed in order to extract useful information. After processing the calibration sequences for each experiment, the validity of each experiment was evaluated. In particular, we assess

Table 1: Experiment duration

Experiment	Frames	Duration [sec]
FF 1	12245	816
FF 2	10983	732
FF 3	1450	96
FF 4	12964	864
FF 5	12029	801
FF 6	16719	1114

Table 2: Calibration agreement

Experiment	RMSE [pix]
FF 1	16
FF 2	9
FF 3	10
FF 4	5
FF 5	12
FF 6	7

the accuracy of the estimated PORs, computed using the first calibration, with respect to the true observed point (the center of the calibration pattern) in the second calibration sequence. The RMS errors for the six experiments are summarized in Table 2, referring to a 640×480 frame.

This preliminary examination is necessary as different factors can degrade the results. The most probable cause in such harsh environments, is that the device might be shifted on the head of the subject due to impacts. The results in Table 2 suggest that RMSE between the calibrations is sufficiently small, hence the data collected in all the experiments are valid. The calibration process, besides allowing estimation of the PORs, it also provides the variance around the estimated point. Using this information we are able to compute a 95% confidence region (CR) around the POR estimation. By processing the data of each experiment, we extract fifteen PORs for each frame of the stereopair, with the corresponding confidence regions.

For each experiment, we recompose the frames acquired by the GM into a video and we overlay the POR and CR information. Additionally the audio captured during the experiments is synchronized with the video sequence. Finally, as the firefighters were talking in their native language, italian, subtitles are added to the final video sequence. These annotated video sequences have been uploaded on the Media server of the TRADR project and are available to the members of the consortium for further analysis and discussion. Example frames taken from the annotated sequences are displayed in Fig. 3.

Finally, local reconstructions of particular parts of the site have been obtained from the gaze machine data. These reconstructions offer a preliminary evaluation of the functionalities that the GM can provide in such scenarios.

Figure 3: Frames taken from the annotated sequences

For example, information regarding structural assessment of the site can be extracted by such reconstructions, obtained from preliminary site inspections. Examples of the 3D reconstructions obtained are shown in Fig. 4.

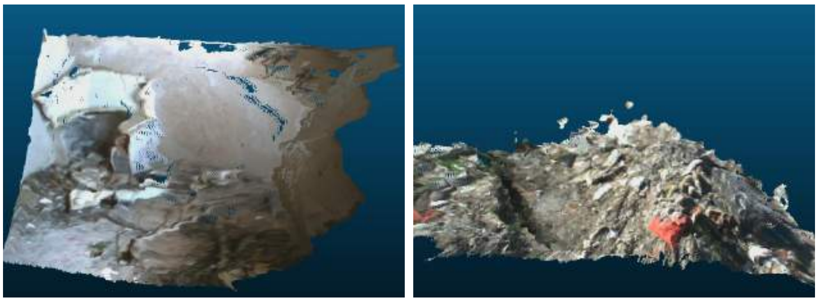

Figure 4: Local reconstructions of scene parts

5 Conclusion

The experiments performed with the GM at Pisa during the last 3 days of TJEx 2014, have helped us obtain useful data regarding the behaviour of the firefighters when they enter hazard areas for inspection and situation assessment. The data collected have been processed and made available to all the members of the consortium for further analysis and elaboration via the Media server. The discussions with the participants have highlighted various aspects which could help in introducing the GM as a tool capable to assist the firefighters in performing their demanding tasks. Moreover, the obtained reconstructions pave the way for the use of the GM for obtaining complete 3D maps of the environment from the data acquired during inspection or other compatible tasks. Finally, a detailed analysis of the acquired data will help the members of the TRADR consortium to obtain a better understanding of the most crucial elements for structural and situation assessment. This can lead to the introduction of different contributions regarding the refinement of the use cases in the TRADR project, the planning of the sorties, as well as the implementation of multi-robot collaboration strategies.

References

[1] F. Pirri, M. Pizzoli, and A. Rudi. A general method for the point of regard estimation in 3d space. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 921–928. IEEE, 2011.

Appendix A Interviews with participants

As discussed in the main text, at the end of each experiment an interview with the participant was conducted. The questions addressed during the interview are the following (translated from italian)

1. Was the run successful according to you?
2. Did it feel artificial performing the task (site inspection) wearing the GM?
3. Was the GM distracting you in any way from performing the task?
4. Did you feel any discomfort because of the GM?
5. Would you change anything in the way the task was performed?
6. Do you have any other comments or suggestions for us?

The responses of the firefighters (FF) are given below (translated from italian). These observations were also summarized and reported at the discussion groups at end of the demo sessions by the participating firefighters.

FF 1 (Day 4 - 1st run)

1. Yes.
2. No, not really.
3. No.
4. No, it was ok.
5. No.
6. No, I don't think so.

FF 2 (Day 4 - 2nd run)

1. Yes.
2. No, it felt natural.
3. No.
4. No, it was comfortable.
5. No, I think the task is fine as it is.
6. No.

FF 3 (Day 5 - 1st run)

1. I think yes.
2. Not at all.
3. No.
4. There was no problem.
5. No, the task is clear.
6. The GM must be more protected in order to survive such harsh conditions. For example it has to conform to some standards typical for FF equipment. It also has to be more compact, for example it should fit approximately where the FF helmet fits. Or for example where the speleological equipment fits.

FF 4 (Day 5 - 2nd run)

1. Yes, all went well.
2. You understand that there is something on your head at the beginning but after a while you focus on the task.
3. No.
4. No, but for long missions it may become more uncomfortable as time passes, e.g. after 30 min.
5. Having a real conversation with another person would make the experiment feel more natural.
6. It seems that the GM is a delicate device, it must conform to the norms, e.g. IP 65, of FF equipment in order to be used in real situations. In particular it has to be more robust.

FF 5 (Day 5 - 3rd run)

1. Yes it was successful.
2. No discomfort at all.
3. I could see the camera in front of the eye but I forgot it as soon as I started the task.
4. Not really.
5. No.
6. It would be good to put some light on the device in order to illuminate dark areas, like speleologists. You have to take care that the light conditions change a lot in these situations.

FF 6 (Day 6 - 1st run)

1. Yes.
2. No everything felt normal.
3. No distraction noticed.
4. It is comfortable.
5. No, the task is good.
6. The GM must be more compact so the firefighters can use it in real situations. It also has to be more sturdy and the weight of the batteries and other supporting equipment must be less to be able to use it in tasks that last longer. It is very important to transmit the images in real-time to the command post so they can take decisions.

DR 7.1: TRADR training and test areas

Joachim de Greeff [*], Salvatore Corrao [†], Norbert Pahlke [‡], Guido van den Broek Humphrey [§]

[*]*TUD, Delft, The Netherlands*
[†]*IFR, Italy*
[‡]*FDDo, Dortmund, Germany*
[§]*GB, Rotterdam, The Netherlands*
⟨J.deGreeff@tudelft.nl⟩

In this document we provide an overview of the three areas (proposed by the three end-users VdF, IFR and RDM) for training and testing of the TRADR scenarios. The yearly TRADR Joint Exercises (T-JEx) are being held on these locations. T-JEx Y1 was done at the area provided by VdF: an old military hospital in Calambrone, Italy.

1 VdF: Calambrone old military hospital

This is a training site of the Corpo Nazionale Vigili del Fuoco (VdF) in Calambrone, Italy, which consist of the ruins of an old military hospital. This site was chosen for T-JEx Y1. Characteristics include the following:

- restricted area
- "hills", ramps
- challenging floors
- small and wide ways
- bright and dark areas
- debris

For an impression, see Figures 1 and 2.

2 FDDO/IFR: "Phoenix" blast furnace ruin

This is a furnace ruin near Dortmund, Germany, which includes a lot of metal; as such posing a considerable challenge wireless networking. Characteristics include the following:

- restricted area
- open and roofed areas within and around the building
- stairs, ramps, "hills"
- challenging floors
- small and wide ways
- bright and dark areas
- complex structure for different levels of difficulty
- platforms for overview during an operation

For an impression, see Figures 3 to 5.

(a) Calambrone 1

(b) Calambrone 2

Figure 1: Impressions of the old military hospital, Calambrone, Italy.

(a) Calambrone 3

(b) The T-JEx Y1 TRADR team photographed from a UAV

Figure 2: Impressions of the old military hospital, Calambrone, Italy.

(a) Phoenix 1

(b) Phoenix 2

(c) Phoenix 3

(d) Phoenix 4

Figure 3: Impressions of the “Phoenix” blast furnace relict, Dortmund, Germany.

(a) Phoenix 5

(b) Phoenix 6

(c) Phoenix 7

(d) Phoenix 8

Figure 4: Impressions of the "Phoenix" blast furnace relict, Dortmund, Germany.

(a) Phoenix 9

(b) Phoenix 10

(c) Phoenix 11

(d) Phoenix 12

Figure 5: Impressions of the “Phoenix” blast furnace relict, Dortmund, Germany.

3 RDM: Industrial training center

Unified Fire Department Services Rotterdam (RDM) have an industrial training center for training and education purposes. Characteristics include the following:

- restricted area
- different types of industry
- stairs, ramps
- challenging floors
- small and wide ways
- bright and dark areas
- pipelines
- platforms for overview during an operation

For an impression, see Figure 6.

(a) RDM 1

(b) RDM 2

(c) RDM 3

Figure 6: Impressions of the RDM industrial training center, Rotterdam, Netherlands.