Usage of Autonomy Features in USAR Human-Robot Teams

Benoit Larochelle *benoit.larochelle@dfki.de*
Language Technology Lab
German Research Center for Artificial Intelligence (DFKI GmbH)
Saarbrücken, 66121, Germany

Geert-Jan M. Kruijff *gj@dfki.de*
Language Technology Lab
German Research Center for Artificial Intelligence (DFKI GmbH)
Saarbrücken, 66121, Germany

Jurriaan van Diggelen *jurriaan.vandiggelen@tno.nl*
TNO Human Factors
Netherlands Organisation for Applied Scientific Research (TNO)
Soesterberg, NL-3769 DE,The Netherlands

Abstract

This paper presents the results of a high-fidelity urban search and rescue (USAR) simulation at a firefighting training site. The NIFTi was system used, which consisted of a semi-autonomous ground robot, a remote-controlled flying robot, a multiview multimodal operator control unit (OCU), and a tactical-level system for mission planning. From a remote command post, firefighters could interact with the robots through the OCU and with a rescue team in person and via radio. They participated in 40-minute reconnaissance missions and showed that highly autonomous features are not easily accepted in the socio-technological context. In fact, the operators drove three times more manually than with any level of autonomy.The paper identifies several factors, such reliability, trust, and transparency that require improvement if end-users are to delegate control to the robots, irrespective of how capable the robots are in such missions.

Keywords:autonomy, transparency, trust, situation awareness, UGV.

1. INTRODUCTION

In NIFTi we investigate how to develop cognitive robots that work together with humans. We consider robots to be, at least to some degree, autonomous actors. If they are not, we should strictly speaking not consider them as team members, but simply as tools. Lackey et al. [1] call it a “shifting paradigm of HRI from a controller/controlled relationship to a cooperative teammate relationship.” We focus on the domain of Urban Search & Rescue (USAR), and particularly where robots support humans early on in making a situational assessment of the disaster site. These missions are physically and mentally stressful, which leads to real-life problems such as misunderstandings, cognitive overload, communication drop-outs, and collisions. Autonomous navigation can thus play a key role improving mission success by lowering the operators' cognitive load and allowing them to focus on other tasks.

However, a robot's autonomous capabilities and intelligence are useless if the humans in the team do not accept the robot as a team member. Recent experiences in a simulated Mars planet (desert) [2] and in the Fukushima earth quake (S. Tadokoro, p.c.) have demonstrated that whenever operators are uncertain what to expect from the robot, or do not trust the autonomy [3], they are unlikely to delegate the control and rather revert to manual control, irrespective of what the robot is able to autonomously perform.

We have jointly developed, with firefighters from the Italian fire brigade (VVF) and the Dortmund fire brigade in Germany (FDDO), a multimodal OCU for a human-robot team with various levels of autonomy[4]. The complete operator control environment allows the operator, via the OCU, to interact with a semi-autonomous unmanned ground rover (UGV), to see the feedback from a teleoperated unmanned aerial vehicle (UAV), and also to use a tactical-level system for mission planning (TRex).

In order to test our robotic system, we recreated a high-fidelity USAR scenario for a human-robot team. While traveling through a tunnel, a truck lost its load of barrels, pallets, and other assorted building materials. This caused a multicar accident where some victims are still trapped in or around cars. Most of the rescue team was in a command post at a remote location where they could safely operate the robot. We investigated how they used the robots, especially concerning autonomous features.Figure 1 shows this end-user evaluation at the *Scuola di FormazioneOperativa* (SFO) in Montelibretti, Italy, a training ground of the VVF.

FIGURE 1: End-user evaluation: tunnel accident scenario with UGV and UAV.

Overview: Below we gather various studies on the use of autonomy and then describe our end-user evaluation. Next, we present various results about operators' activities, focusing mostly on the use of autonomy features. Finally, we discuss causes and possible improvements for the acceptance of these features.

2. BACKGROUND

As early as 2004, Burke *et al.*[5] postulated that operators need adequate awareness of the robot's state and surroundings if they are to release control and use the robot's autonomy. Their suggestions have been since demonstrated in several different contexts.

In fact, improving situation awareness has always considered as a highly important issue in USAR robotics. For example, Yanco& Drury [6] performed a study of operator performance at the AAAI Robot Rescue Competition in 2002, 2003, and 2004. The authors highlighted the importance of large video feeds and the integration of all necessary information and controls in a single window. Otherwise, operators have more difficulty integrating the robot's perspective into their mental map of the area [1], [5], [6].During the three years of competition, several robots had autonomous functionalities, but most of these features were not used as the teams preferred to manually control the robots.

The recent incidents at the Fukushima Daiichi nuclear disaster site also exemplified the lack of acceptance of autonomy, as discussed by S. Tadokoro at the 2011 AAAI Fall Symposia. With low situation awareness and difficult terrain and obstacles, the operators brought a second robot only to see the main one from an exocentric perspective. The operators also preferred to manually control the robots in such difficult situations. [7]

Finally, autonomy acceptance problems exist also in the context of asynchronous interaction. Stubbs *et al.*[2] presents an outdoor robot with various levels of autonomy to explore a simulated Mars planet (desert). The authors explain that as the robot's autonomy increases, the traditional problems of perception and situation awareness leave place to problems of transparency and trust. Their conclusion is that robots must adapt their behaviors to create more realistic conversations with users. Comparing, and ideally completing, each other's knowledge should help in achieving more common ground and transparency, which are necessary if the operators are to accept – and use – the robot's autonomy.

However, sharing knowledge to establish common ground and shared situation awareness is a daunting task. Many parameters come into play, such as the users' skill levels, their familiarity with the task and the environment, the task itself, the type and modality of the information, and the timing and frequency of the exchanges between the operator and the robot [1], [8].For example, Parasuraman*et al.*[8] discuss adaptivity in providing information to operators, as well as how to avoid pitfalls of shared initiative systems. Lackey *et al.*[1] discuss how different sources of information must be understood as a whole to create high-level situation awareness in the context of mixed-initiative soldier-robot teams. The authors also prone "sharing information back-and-forth in a fluid natural manner using combinations of communication methods." Torrey *et al.*[9] shows that when executing a robot-guided task, the robot under- or over-specifying objects to which it refers can lead not only to performance problems but also to a degradation of the social cohesion. The paper also demonstrates that this phenomenon is amplified under time pressure.

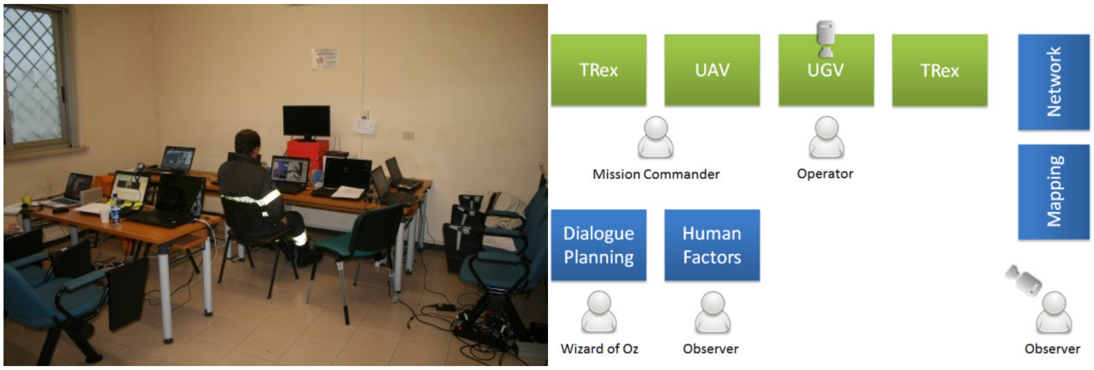
3. END-USER EVALUATION

In the NIFTi project, the requirements, design, and testing phases have all been performed jointly with firefighters (end-users) from the Italian fire brigade (VVF) and the Dortmund fire brigade in Germany (FDDO). This collaboration allowed us to create highly realistic scenarios and systems, and to test them directly with end-users.

Location and Setup

In December 2011, we recreated a tunnel car accident at the SFO training site, shown in Figure 1. The area spanned 25 meters into the tunnel by a width of 10 meters, filled with debris, pallets, barrels, crashed vehicles, and smoke. Figure 5 shows a map of the area, where each grid cell represents 1 m^2. Participants had to assess the situation with one UGV and one UAV in 30 or 40 minutes, depending on whether the autonomy features were activated. The users received 30 minutes of training with the OCU, plus 15 minutes for the autonomy features. They also performed a few navigation tests before starting the scenario [10]. Ten participants, one each morning and one each afternoon for a week, participated in the experiment out of which we analyzed six complete data sets. The other four time slots were incomplete due to technical or logistical problems.

The scenario consisted of a team of responders: in the field, aUAV pilot; in a remote command post, shown in Figure 2,a mission commander and a UGV operator (experimental subject).The front row consisted of the computers that the firefighters could access. One computer with TRex[11]was available for each one of them, an OCU connected to the UAV was placed between the two, and an OCU connected to the UGV was directly in front of the operator. This set-up allowed both the operator and the mission commander to have access the high-level features of TRex, while being able to zoom in to the local situation awareness provided by any of the robots through the OCUs.



a) Picture of the command post b) Diagram of the command post

FIGURE 2: SFO December 2011: Command Post.

The other computers were used for support and debugging. One acted as a DNS/NTP server, one ran the mapping algorithms,one ran the dialogue and planning components, and finally one collected data about human factors such as heart rateand emotion through a facial recognition software. Instead of an automaticspeech recognizer,we opted for a Wizard of Oz approach, which eliminated problems due to noise and poor language skills in English. Two observers and two cameras were used.The first was a webcam clipped on the main OCU and thesecond was a standard video camera on a tripod, capturing abroad view of the scene.The set-up allowed also all support staffto oversee the experiment and freely work without disturbingthe participants.

The NIFTi System

The NIFTi system is composed of several components. First, the UGV consists of a man-portable robot with passively adaptable left and right tracks, each with motorized flippers at the front and back[12]. It has an omnidirectional camera and a rotating laser. A man-portable micro-copter was also developed to provide video feeds from two cameras.Because end-users are not accustomed to using robots and since they will be using the system under difficult conditions (i.e. varying cognitive load, high stress, loud environment, time pressure, etc.), interaction paradigms with these robots must be natural and intuitive. The UAV was actually maneuvered only from a trained pilot who received instructions from the mission commander. The video feed was broadcast in the command post.

The OCU [4]is multimodal because its two main modes of input are voice and touch with the laptop's built-in microphone and 15.6" dual-touch screen. In addition to displaying the robots' cameras' video feeds, a virtual scene is available, showing a map built up as the robot explores the environment.Laser points representing the obstacles in front of the robot and a 3D robot model are also shown. It is possible to overlay the virtual scene on top of a camera feed, which helps operators navigate in low or varying visibility (e.g. darkness, smoke).The OCU can display one, two, or four of these views simultaneously, as shown in Figure 3.The robot can automatically detect cars and victims and tells the operator via speech and text, in addition to placing iconsin the virtual scene.The robot can be manually navigated with the touch screen, but it also understands vocal commands, such as “Move forward”, “Turn right” and “Go to the car”.

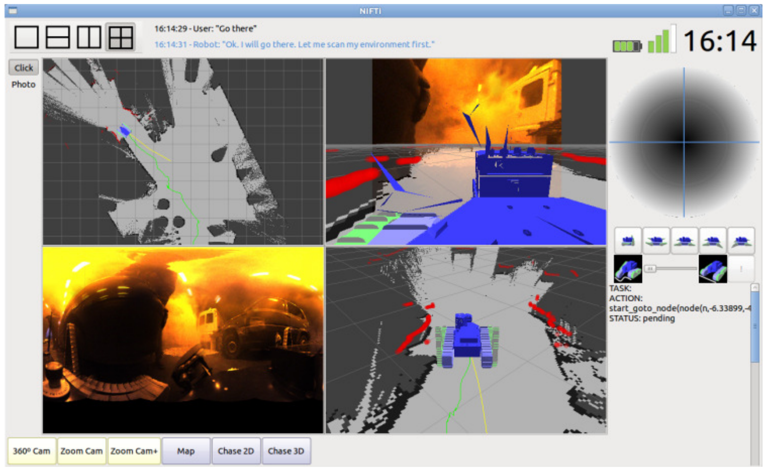


FIGURE 3: The NIFTi OCU in the tunnel accident scenario.

In addition to the OCU, we wanted to provide the rescue team with a higher-level system, hence the TRex computers. In USAR scenario, we consider that robots are operated at three levels, as detailed in [13]:

- Executional: low-level, short elementary actions (e.g. accelerating, observing objects)
- Operational: mid-level, executing a plan of actions (e.g. following a route defined during the mission)
- Tactical: high-level, planning the resources and steps (e.g. which robots will investigate which areas)

The OCU supports the executional and operational levels while the TRex system [11]supports the tactical level. The positions of the UGV and the UAV can be visualized in the TRex map, as well as localized icons representing pictures taken by the operator through the OCU and reports added by the operator and the mission commander.

4. RESULTS

We collected data for six successful missions, three with autonomy features, and three without. We are aware than this data set is quite small, but we chose to create a high-fidelity simulation with real firefighters instead of a typical lab experiment with students, even if it meant reducing the number of participants. The availability of the site and of end users made it impossible to extend the experiment to more than one week. Thus, this paper does not claim to have statistical significance like many indoors robotics experiments, but presents more data and analysis than a field report.

We synchronized the two video streams fromthe observers with recordings from all computers at the command post to prepare the results presented in this section. We expected to see clear changes in biophysical data during the missions, patterns in human-human and human-robot communication, similar driving and exploration styles among the firefighters, as well as enthusiastic use of autonomy. The results were quite different than what we expected and the salient points are presented below.

Operators' Activities

Figure 4 presents the time distribution of the six users during the scenario[1]The diagrams show that the operators spent on average 57 % of their time navigating, but with high variability. These results are very similar to last year's end-user evaluation [15] with an average of 54 % (varied from 47 % to 62 %). Burke et al [5] showed a slightly lower figure, 44%, but also with great variability. They mentioned, however, that “operators spent significantly more time gathering information about the state of the robot and the state of the environment than they did navigating” and that they “had difficulty integrating the robot's view into their understanding of the search and rescue site. They compensated for this lack of situation awareness by communicating with team members at the site”. We have also experienced these problems in another field trial in July 2011 [14], but not in this end-user evaluation.

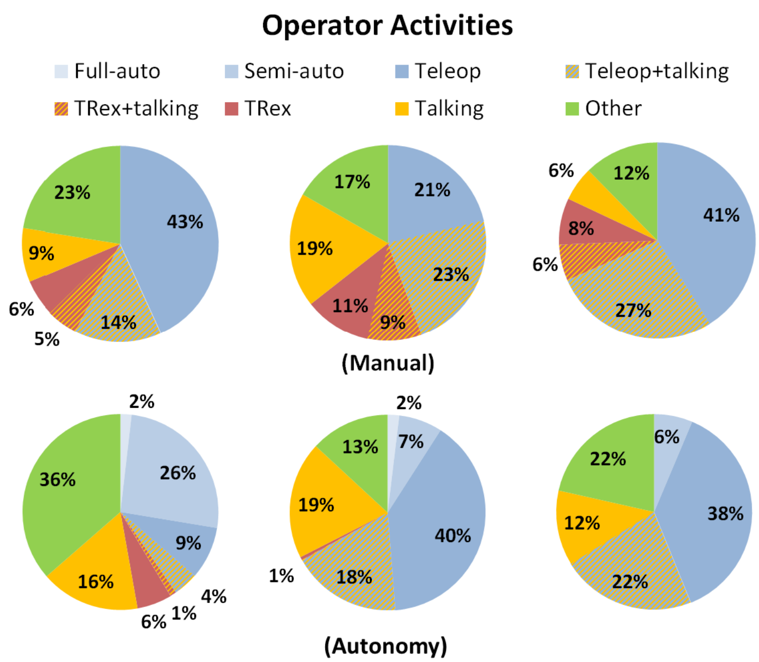


FIGURE 4: Operator activities for the six participants during the scenario.

We can also notice that the operators spent on average 35 % of their time speaking with the mission commander, again with high variability ranging from 21 % to 51 %. While talking, the operators were also navigating the robot on average 50 % of that time, with a range from 19 % to 70 %. This indicates that the operators' cognitive load was not too high to perform these two simultaneous tasks, contrary to our expectations.Finally, we notice that the operators spent on average 20 % of their time on other activities, mostly moving the camera and studying the environment through the image.

In addition to overall time distribution, we analyzed *when*did the operators do*what*. Once again, we could not identify any recurring pattern, but rather observed varied styles again. However, an interesting observation is that the users switch to a lower autonomy mode mostly after a failure of the autonomous feature. Desai *et al.*[3] also observed that users switched to lower autonomy modes quickly after the robot made mistakes, and took much longer to trust the autonomy again.Finally, the participants' heart rates were monitored but they showed nearly no variation

[1]Because we were not able to get a firefighter available for the whole week to play the role of the mission commander, we had one for the first three participants and another one for the last three participants. Their styles of interaction partly explain the large difference in the use of TRex.

during the scenario. They were also asked to indicate their cognitive load on a scale of 1 to 5 every two minutes, and no significant variation was observed.

Operators' Paths and Performance

Our expectations were that the operators would generally navigate around areas of interest, based on their firefighting training. However, and analysis of their paths revealed to be highly varied and we could not extract patterns in driving styles or prominent locations for stopping and observing the scenario. Figure 5a) shows the path of one of the participants, augmented with spheres that indicate how much time she spent at each location. The smallest spheres indicate 1 second while the largest indicate 15 seconds or more. Additionally, arrows with numbers indicate where the robot was at every two-minute interval.

Figure 5b) shows the same path but color-coded to indicate the level of autonomy used. The green sections show where the robot was teleoperated. The orange sections shows that the user was using semi-autonomy (short commands such as “Go forward” and “Turn left”).All users started with operating the robot under autonomous mode. However, they all took back control as soon as the path became more difficult to navigate and several objects to inspect became visible. Some of the operators used autonomy features again later in the missions, but only for small movements not visible on the map. When the robot asked the users if it should autonomously go to a newly detected car, they ignored the question and continued teleoperating.

Contrary to the high variability in the paths of the six participants, their performance in finding scenario elements were quite similar. Cars were always reported, victims were found 79 % of the time, and danger signs 44 %. The results also indicate that no element was particularly hard to find and that no difference exists between the participants with and those without autonomy features.

Collisions and Situation Awareness

Teleoperation is usually considered ‘bad’ because it leads to frequent collisions. Our results are comparable to other studies. Table 1 shows collision data from the NIST competitions [6] and from the first NIFTi end-user evaluation [15]. Unfortunately, it is not possible to directly compare the numbers because too many variables are present. For example, scenario sizes and densities, time pressure, robot platforms, and OCUs influenced the number of collisions. In addition, we provided little user training on the NIFTi platforms but the pilots in the NIST competitions were well trained developers of the systems.

Event	NIST 2002	NIST 2003	NIST 2004	NIFTi Jan. 2011	NIFTI Jan. 2011	NIFTI Dec. 2011
Robot	Various	Various	Various	Generaal	P3-AT	NIFTi
Duration	Max 20 min.	Max 20 min.	Max 20 min.	15 min.	Max 15 min.	30, 40 min.
Collisions	6.2	2.2	1.3	3.2	1.3	9.2

TABLE 1: Collisions in the scenario.

Usage of the OCU Views

The OCU was always launched with four default views (shown in Figure 3), and all operators except one used them without any modification. In fact,some of the operators did not even use all of the views.More specifically, the ‘Map’ view, which shows an overallpicture of the scenario,was not used by all operators. Moreover, the mission commanders used the map view while the operators were looking at a different part of the screen. This is an interesting behavior, since both users were given a computer with TRex, which has more high-level functionalities than the OCU, but they often converged to using a single laptop. Similarly, it was observed in the NIST competitions [6] that the screens other than the main one often get ignored. In our case, it was also easier for the participants to integrate the robot's perspective with the map view than with the TRex system.

FIGURE 5: Path Followed by One of the Participants.

5. DISCUSSIONS

Considering the expectations that we had about the use of autonomy, we can certainly say that the results are disappointing. Figure 4 and Figure 5b) clearly show how little autonomy features were used. More precisely, the operators drove three times more manually than in all autonomous modes combined. Despite these results, we continue to believe, based on studies such as [3], [6], that more autonomy would benefit the users; either in the number of collisions or in victim discovery performance. We thus present here problematic areas of the NIFTi system and evaluation methodology that impacted the use of autonomy.

Technical Reliability & Flexibility

The NIFTi platform was produced in 2011 and being inexperienced, we set the safety margins too high. In consequence, the robot often stayed still rather than risking navigating near objects or into unknown space. Since the goal of the mission was to explore space, the users quickly got frustrated and switched to a lower autonomy mode. Short commands (e.g. “Move forward”) worked well, but did not offer the same flexibility as manual control. The operators sometimes wanted to go at a specific place, so they did it themselves. To solve that problem, we are developing an approach to analyze a robot's surroundings to provide a functional-geometric interpretation of movement commands such as "Move forward". A correct interpretation will allow the robot to move an appropriate distance based on the environment rather than moving a fixed amount. We will run experiments to determine if such a behavior leads users to rely more on autonomy[16].Finally, the robot's autonomous modes were very slow. By comparing Figure 5a) and b), we see that the first stretch took more than four minutes, at which point the user started manually driving. Autonomy was never used late in the scenario when the time pressure was higher.

Cognitive Load

One of the goals of autonomous features being the reduction in cognitive load, the features are most useful under high load. However, our users indicated that their cognitive loads were always moderate – this corroborates with them talking while teleoperating. Questionnaires also showed that they did not consider teleoperation or the mission in general to be very difficult. They had thus little incentive to use any autonomy. In addition, Oviatt*et al.*[17] found out that users interact in a multimodal fashion mostly when the task at hand is difficult and the information to convey is complex. In our case, the operators controlled the robot – in a unimodal way – because it was easy and not hindering their other tasks. Gómez [18] also ran an experiment that points to the same conclusion. In his case, operators controlled either one, two, or three robots. Operators teleoperated the single robot 93 % of the total navigation time, compared to 48 % with two robots and 27 % with three.

Engaging Dialogue

Contrary to our expectations, the users never got engaged in a true dialogue with the robot. Since the robot was silent most of the time, except when detecting cars or responding to spoken navigation commands, the users did not feel that the robot was talking to them, but rather was giving debugging information. One problem is that the spoken information contained spatial information, which was not presented to the user. For example, when detecting cars, the robot alerted the operators, but did not show where they were located on the map. When prompted to “go to the car”, the operators simply ignored the question.Comparatively, Torrey *et al.*[9] showed that when executing a robot-guided task, the robot under- or over-specifying objects to which it refers can lead not only to performance problems but also to a degradation of the social cohesion. The paper also demonstrates that this phenomenon is amplified under time pressure.

Transparency

From past experiences, we believed that reducing the needto teleoperate the robot would free up some time for the userto observe the environment or perform other tasks. However,autonomous robot behavior must be transparentto the operators;otherwise, they will not understand itand will be unlikely to relinquish control to the robot. Withouttransparency, not enough trust is built up and the robot remainslargely teleoperated by the operators. In fact, the negative impacts of low transparency on human-robot interaction have been suggestedbefore in [5] and were observed in [2]. More recently, S.Tadokoro discussed the same problems at the 2011 AAAIFall Symposia about experiences at the Fukushima accidentsite. Our end-user evaluation confirmed theseobservations.For example, we ensured that the NIFTi robot would always give feedback when it succeeded or failed a task, but it never explained why it failed. Given that the users did not know about the robot's safety margins, they were left confused about the robot's autonomous behavior and wondering what happened. In successful cases, the planned path was not displayed (due to technical reasons), which also made users nervous about letting the robot autonomously navigate. In many cases, the operators were wondering if they should stop, wait or try something again.

Trust and Expectations

Our users received training for manual control, in which they usually did not crash, as well as for autonomous control, in which the robot crashed a few times. These events could have led them to trust in their abilities more than in the robot's autonomy. Desai *et al.*[3] showed that in such cases, users tend to use manual control. The users, unfamiliar with robots, also expected more reliability and functionality. In particular, they expected the flippers to automatically adjust, regardless of the autonomy mode. Komatsu and Yamada [19] showed that when agents' functionalities are lower than the users' expectations, users tend to stop interacting with these agents. Analogously, our operators stopped using autonomy features after having tried them and being disappointed.

6. FUTURE WORKS

While the NIFTi project continues to work on autonomy features, it also aims at improving the human-robot interaction during teleoperation. Questionnaires about the OCU showed that the operators did not complain about anythingparticularly bad in the OCU. They preferred manual drivingin certain cases, automated in others. Unfortunately, they did not identify what classes of scenarios or environments prompt manual override. In any case, we expect that more operating experience would be required to make such judgments.

The main request from the operators wasto improve the display of distances. Since all users made5 to 20 collisions in the scenario, we consider that an improvement is required.After the evaluation, we decided to superimpose concentric circles at 1, 2, and 5 meters around the robot's 3D model. With these aids, it is much easier to estimate distances to the surrounding obstacles. We have also added a telescopic arm and are working on a new virtual camera, both of which allow raising the point of view of the cameras and hence projecting better depth perception. The traveled path is also now shown by default.

The next problem is that even with these improved views,it is not guaranteed that the operators will use them more. Automatic adjustmentof the views was not implemented because wefirst wanted to collect data on how the operators used them. Given low usage results, we need to find innovative ways to adjust the views for the users. Onesuggestion is an automatic zoom, which zooms in on the robotat low speeds or when navigating close to obstacles. Thesensitivity of the control widget could also be adjusted withthese parameters. Such features are already available in carsand embedded navigation systems. In a subsequent end-user evaluation in November 2012, we ran a cognitive model during the missions in order to evaluate the cognitive load of the operators. Once we analyze the results and determine that they correlate with reality, we will investigate how to adjust the views in a non-disruptive manner.

Alternatively, we are also working on the integration of in-field pictures taken from either the UGV, the UAV, or an in-field rescuer. All of the pictures will be centrally collected and stored at the command post, and made available in the OCU and in TRex. Because these pictures will be geo-located, we will show them as icons on the maps, and the operators will be able to see the environment from different points of view, helping with navigation and situation awareness in general.

7. CONCLUSIONS

We organized an end-user evaluation and recorded data from six tunnel car accident missions. Firefighters used the NIFTi robots and OCU as part of a human-robot team. We have observed highly varied usage patterns, with respect to exploration strategies, driving styles, and use of autonomy. Users spent 57 % of their time navigating, although mostly manually. Autonomy features were not extensively used, and switch to lower autonomy modes happened mostly after autonomy failures.We identified several factors that could have led to low usage of autonomy and discussed several improvements that we are developing. In particular, transparency is needed for trust, and trust is needed for autonomy. Thus, the robot should be more communicative and transparent about its status and actions, other robots, and the environment. Statements that carry spatial information should convey this spatial part in a multimodal fashion. We would also like to spend more time on user training, allowing them to adjust their expectations and develop trust in the system. Additionally, we need more focused experiments to separate the effects of technical limitations versus those social effects onto the usage of autonomy features.

8. ACKNOWLEDGEMENTS

The authors would like to thank the end-user organizations, VVF and FDDO, for their support. They have been very accommodating in letting us use their infrastructures and have provided numerous hours of their time to help us collect valuable data and feedback.

The research described here has been financed by the European Integrated Project NIFTi under contract number FP7 ICT 247870, Cognitive Systems and Robotics.

9. REFERENCES

[1] S. Lackey, D. Barber, L. Reinerman, N. I. Badler, and I. Hudson,“Defining next-generation multi-modal communication in human robotinteraction,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55. SAGE, September 2011, pp. 461–464.

[2] K. Stubbs, P. Hinds, and D. Wettergreen, “Autonomy and commonground in human-robot interaction: A field study,” *IEEE IntelligentSystems*, vol. 22, pp. 42–50, March 2007, special Issue on Interactingwith Autonomy.

[3] M. Desai, M. Medvedev, M. Vázquez, S. McSheehy, S. Gadea-Omelchenko, C. Bruggeman, A. Steinfeld, and H. Yanco, “Effects ofchanging reliability on trust of robot systems,” in *Proceedings of theseventh annual ACM/IEEE international conference on Human-RobotInteraction*. ACM, 2012, pp. 73–80.

[4] B. Larochelle andG.J.M. Kruijff, “Multi-view operator control unit to improve situation awareness in USAR missions,” in*Proceedings of the 21th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 1103–1108.

[5] J. Burke, R. Murphy, M. Coovert, and D. Riddle, “Moonlight inMiami: An ethnographic study of human-robot interaction in USAR,”*Human Computer Interaction*, vol. 19, no. (1–2), pp. 85–116, 2004.

[6] H. Yanco and J. Drury, “Rescuing interfaces: A multi-year studyof human-robot interaction at the AAAI robot rescue competition,”*Autonomous Robots*, vol. 22, no. 4, pp. 333–352, May 2007.

[7] E. Guizzo, E. Ackerman, M. Waibel, M. Taylor, andS. Bouchard, “Fukushima robot operator writes tell-all blog,”http://spectrum.ieee.org/automaton/robotics/industrial-robots/fukushima-robot-operator-diaries, Aug. 2011, contains the Englishtranslation of the Japanese blog from the anonymous author S.H.

[8] R. Parasuraman, K. Cosenzo, and E. de Visser, “Adaptive automationfor human supervision of multiple uninhabited vehicles: Effects onchange detection, situation awareness, and mental workload,” *MilitaryPsychology*, vol. 21, pp. 270–297, 2009.

[9] C. Torrey, A. Powers, M. Marge, S. R. Fussell, and S. Kiesler,“Effects of adaptive robot dialogue on information exchange and socialrelations,” in *Proceedings of the 2006 ACM Conference on Human-Robot Interaction*. Press, 2006, pp. 126–133.

[10] T. Mioch, N. Smets, and M. Neerincx, “Assessing human-robotperformances in complex situations by means of unit task tests,” in*Proceedings of the 21th IEEE International Symposium on Robot andHuman Interactive Communication*. IEEE, 2012, pp. 621–626.

[11] J. van Diggelen, K. van Drimmelen, A. Heuvelink, P. Kerbusch,M. Neerincx, S. van Trijp, E. Ubink, and B. van der Vecht, “Mutualempowerment in mobile soldier support,” *Special Issue of the International Journal of Battlefield Technology on Human Factors and Battlefield Technologies*, 2012.

[12] Bluebotics, “Patent, mobile robot,” Patent, 2011, pCT/EP2011/060937.

[13] J. van Diggelen, R. Looije, T. Mioch, M. A. Neerincx, and N. J. M.Smets, “A usage-centered evaluation methodology for unmannedground vehicles,” in *Proceedings of The Fifth International Conferenceon Advances in Computer-Human Interactions (ACHI)*, 2012.

[14] G.J.M. Kruijff. et al, “Experience in system design for human-robot teaming inurban search & rescue,” in *Proceedings of Field and Service Robotics(FSR) 2012*, Matsushima/Sendai, Japan, 2012.

[15] B. Larochelle, G.J.M. Kruijff, N. Smets, T. Mioch, and P. Groenewegen,“Establishing human situation awareness using a multi-modal operatorcontrol unit in an urban search & rescue human-robot team,” in*Proceedings of the 20th IEEE International Symposium on Robot andHuman Interactive Communication*. IEEE, 2011.

[16] S. Keshavdas andG.J.M. Kruijff, “Interpretation of Vague Scalar Predicates Expressing Direction.”Technical Report, February 2013.

[17] S. Oviatt, R. Coulston, and R. Lunsford, “When do we interact multimodally?Cognitive load and multimodal communication patterns,” in*Proceedings of the International Conference on Multimodal Interfaces*. ACMPress, 2004, pp. 129–136.

[18] A. V. Gómez, “Evolutionary design of human-robot interfaces forteaming humans and mobile robots in exploration missions,” Ph.D.dissertation, Universidad Politécnica de Madrid, 2010.

[19] T. Komatsu and S. Yamada, “Adaptation gap hypothesis: How differencesbetween users’ expected and perceived agent functions affecttheir subjective impression,” *Journal of Systemics, Cybernetics andInformatics*, vol. 9, no. 1, pp. 67–74, 2011.