# Design and Usability Testing of a Text-Based Simulation for Ethics Training in Disaster Robotics

Linda Battistuzzi
Dept. of Informatics, Bioengineering, Robotics and Systems Engineering
University of Genoa
Genoa, Italy
linda.battistuzzi@unige.it

Lucrezia Grassi
Dept. of Informatics, Bioengineering, Robotics and Systems Engineering
University of Genoa
Genoa, Italy
lucrezia.grassi@edu.unige.it

Carmine Tommaso Recchiuto
Dept. of Informatics, Bioengineering, Robotics and Systems Engineering
University of Genoa
Genoa, Italy
carmine.recchiuto@dibris.unige.it

Antonio Sgorbissa
Dept. of Informatics, Bioengineering, Robotics and Systems Engineering
University of Genoa
Genoa, Italy
antonio.sgorbissa@unige.it

***Abstract* - Rescue robots are expected to become commonplace at disaster sites, where they are increasingly being deployed to provide rescuers with improved access and intervention capabilities while mitigating risks. The presence of robots in operation areas, however, is likely to carry a layer of additional ethical complexity to situations that are already ethically challenging. In addition, limited guidance is available for ethically informed, practical decision-making in real-life disaster settings, and specific ethics training programs are lacking. The contribution of this paper is thus to propose a tool that will support ethics training for rescuers operating with rescue robots. To this end, we have designed an interactive text-based simulation. The simulation was developed in Python, using Tkinter, Python's *de facto* standard GUI. It is designed in accordance with the Case-Based Learning approach, a widely used instructional method that has been found to work well for ethics training. The simulation revolves around a case grounded in ethical themes we identified in previous work on ethical issues in rescue robotics: fairness and discrimination, false or excessive expectations, labor replacement, privacy, responsibility, safety, and trust. Here we present the design of the simulation and the results of usability testing, which included a qualitative exploration and a quantitative survey.**

***Keywords* –** disaster robotics, rescue robotics, ethics training, case-based learning, text-based simulation, usability.

## I. Introduction

Disaster robotics, also termed rescue robotics, is a discipline within field robotics whose goal, according to leading expert Robin Murphy, is to "enable responders and other stakeholders to sense and act at a distance from the site of a disaster or extreme incident." [1]. Robotic technologies deployed at disaster sites can allow rescuers to reach areas that would be too dangerous or slow to access otherwise. In addition, they make it possible for humans to interact with the environment through remote sensing devices. Unmanned Aerial Vehicles (UAVs), for instance, can carry cameras and thermal imagers to supply aerial views of a disaster site, improving situational awareness and helping in the identification of casualties. Some UAV models can also deliver payloads, such as a cellphone, water, food, or first aid supplies [2]. An Unmanned Ground Vehicle can place sensors to monitor the situation or remove debris to help extricate trapped victims [3], [4], while an Unmanned Marine Vehicle can carry out above and below waterline inspection of roads, bridges, and seawalls [5]. Overall, by affording rescuers improved access and intervention capabilities, rescue robots are expected to help save lives, reduce injuries, and speed up recovery [1].

Owing to the growing impact of natural and man-made disasters, the need for rescue robots is projected to increase in the coming years [1]. At the same time, significant advances have been recently made in the field: legged robots are becoming more adept at negotiating rough terrains; improved human–robot interfaces are making for better interactions between operators and robots, and developments in UAVs are leading to new applications for aerial platforms [6] . Some of these advances are already used in commercial products, thus further increasing the likelihood that rescue robots will be widely adopted.

Disasters pose unusual, sometimes unique problems, placing rescuers in unfamiliar situations, and are often characterized by a range of ethical challenges. Many of those challenges are associated with the dangerous, unpredictable, and fast-changing conditions under which responders must operate; others are linked to the lack of resources and time that characterize their work. Faced with situations in which the professional ethics they are guided by in routine emergencies may not apply [7], responders need to make ethically difficult decisions, yet limited guidance is available for ethically informed, practical decision-making in real-life disaster settings, and specific ethics training programs are lacking [8].

Although rescue robots' presence in operation areas is likely to raise additional ethical concerns, the ethical issues associated with rescue robotics have been largely unexplored [9]. In light of this, we recently conducted a scoping review of the ethical issues associated with the deployment of robotic technologies in disaster scenarios [10]. We found that questions associated with fairness and discrimination were the most frequently discussed, along with false or excessive expectations, labor replacement, privacy, responsibility, safety, and trust. Our results confirmed that the literature on ethical issues in rescue robotics is scant and disparate, but the publications we identified in our review uniformly endorsed a proactive approach to their management [10].

The contribution of this paper is therefore to propose a tool that will support ethics training for rescuers who operate or interact with rescue robots. To this end, we have designed an interactive text-based simulation, based on an ethical case grounded in the themes identified by our scoping review. Here we report on two sets of results: the text-based simulation resulting from the design process, and the results of usability testing of the simulation - including a qualitative exploration and a quantitative survey. The article is organized as follows: Section II provides an overview of the case-based approach to ethics training; Section III describes the text-based simulation, and Section IV reports the results of the usability testing of the simulation. We close with Conclusions in Section V.

## II. CASE-BASED ETHICS TRAINING

Ethically challenging situations typically present competing goals, principles or values, so that resolutions can be difficult to reach and are often not univocal [11]. When attempting to make sense of situations like these, reflecting on experiences can be helpful [12]. This type of experiential knowledge can be gained vicariously via case-based learning (CBL), an instructional method that has been widely used across multiple disciplines [13]. CBL presents learners with problematic study cases designed to mirror real-life scenarios, which they are asked to solve by identifying analogies, building inferences, and forming decisions [14], [15]. The stories, characters and scenarios within case-studies are tools through which learners develop new understandings.

CBL is considered to be more effective than traditional lecture-based approaches in improving decision-making, critical thinking and reasoning skills [14], [16]. Moreover, it is often reported to be more enjoyable than lectures, which may have a positive impact on learner engagement [12], [14].

CBL has been extensively applied in ethics education as cases can be designed to replicate the nuances of ethically complex situations [17], [18]. In addition, cases provide learners with opportunities to practice engaging with such situations, [12] but also to think about and discuss abstract ethical concepts [19]. CBL may be especially well-suited for individuals who have little or no experience with managing similar ethical problems because cases allow for the development and practice of skills without having to rely on any actual experience of the circumstances [20].

The pedagogical effectiveness of CBL can be promoted or hampered by case content and structure [21]. It has been suggested that even simple cases should always feature a specific setting, a logical sequence of events, and defined characters [13]. An extensive review on CBL conducted to identify strategies for optimal case construction [14] concluded that, in order to be effective, cases should be relevant, realistic, engaging, challenging, and instructional. The authors of the review also proposed different strategies to integrate these features into cases, including: (i) placing cases within realistic settings that are relevant to learners, (ii) adding depth by providing appropriate information about the characters and issues at hand, and (iii) increasing complexity by adding or concealing certain information [14]. In addition, cases should provide enough detail about crucial aspects to enable critical thinking, but, at the same time, they should not be so long that they become tedious [22].

Features of ethics cases that may interfere with their educational effectiveness have also been described. As pointed out by Beal & Orbison [23], correct responses or the appropriate course of action may be obvious. Learners may view the events described in the case as unlikely, or the scenario presented as unrealistic. Furthermore, cases seldom feature multiple decision points, in which one decision will compound the ethical problem at hand or constitutes itself an ethical problem that will need to be addressed via subsequent choices [23]. In real-world scenarios, complex ethical situations often emerge from the combined effect of multiple decisions with accumulating negative outcomes. Such outcomes are often difficult to predict and may generate decision fatigue [24] so that learners' ability to choose the most ethical conduct deteriorates as they are faced with multiple decisions.

Our intention here was to design a simulation that will actively engage trainees and should be undertaken by teams of two or more, thus requiring discussion and deliberation, and introducing peer pressure influence on decisions [23]. The purpose of the simulation is to enhance trainees' awareness of ethical issues that may arise when deploying robots at disaster sites, as well as their ability to identify those ethical issues and understand the consequences of decisions made to solve or manage them. We envision the simulation as being one element of a more comprehensive ethics training approach that would also include a theoretical component [25].

## III. SIMULATION DESIGN

Drawing on CBL and on the work described by Beal & Orbison [23], we designed a multi-branched, interactive ethics simulation that follows the strategies for case construction and addresses the limitations mentioned earlier. In developing the simulation, we were guided by the following requirements:

- the case presented in the simulation should be realistic, i.e. learners should be able to identify with the situation described and consider the issues presented to be relatively common in real-life disasters;

- the case should require learners to engage in decisions at multiple points, and those decisions should shape the subsequent chain of events;

- the case should elicit decision fatigue;

- the case should involve personal costs;

- the system should prevent trainees from changing their decisions;

- the main character in the case should be gender-neutral and be in a position to make the decisions described in the case [21].

The activity is administered through a computer application developed in Python, using Tkinter, Python's *de facto* standard GUI. The choice fell on Tkinter as it is lightweight, cross-platform, and visual elements are rendered using native operating system elements.

When the application starts, the trainee is welcomed with a screen containing the title of the application and a guide with some useful information about the purpose of the simulation, the actions that can be performed, and the range of the final score. When the "Continue" button at the bottom of the screen is clicked, the GUI displays the first narrative segment along with an associated image, as shown in **Fig. 1**. Once the trainee has finished reading the displayed text, they can click the “Continue” button to proceed: the new text is added below the previous one, and dashes are used to

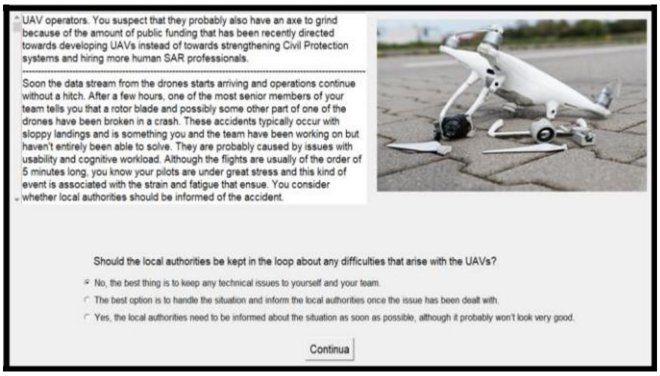


**Fig. 1** Simulation screenshot.

provide visual separation. A scroll bar allows the trainee to go back in the story if needed. If the text is followed by a question, the question and the answers are shown in the bottom part of the screen, below the text and the image. The trainee can choose their answer by selecting the corresponding radio button, and then clicking “Continue” to confirm the choice. A subsequent narrative, specific to the trainee’s decision, is displayed, and another set of decisions is eventually presented. At the end of the story, a completion message is displayed along with the score.

Our case involves a team of University drone pilots and researchers who have been called by the National Civil Protection to provide rescue support at the site of an earthquake. The team has little experience of actual rescue missions but is technically competent and has all the required authorizations and certifications in place to operate during disaster missions. The primary actor in the case is identified as a relatively senior University researcher and drone expert who has participated in meetings with the National Civil Protection and local authorities to discuss the capabilities of drones in disaster settings and establish agreements to cooperate. During these meetings, the character has found that local officials and members of professional rescue organizations are sometimes skeptical about how simple and useful it might be to deploy drones or other robots in real-life disaster scenarios. The character understands that this resistance may have been caused by negative experiences with well-meaning but ill-prepared amateur UAV operators, but also by the amount of public funding that has been recently directed towards developing UAVs instead of, perhaps, towards strengthening Civil Protection systems and hiring or training more human rescue professionals. Encountering this type of resistance increases the character’s feeling of responsibility in terms of proving that the technology is safe and useful, thus generating trust in the technology among potential users, and of advancing their own and their teams’ reputation. Issues of trust are indeed one of the main ethical themes throughout the simulation, along with concerns associated with transparency, fairness, community engagement, hype and excessive or false expectations, and labor replacement.

There are several different paths through the case, with different endpoints reflecting realistic outcomes based on the decisions made, as seen in Figure 2. Each path requires the trainees to make a minimum of four, and as many as eight, decisions in the simulation. Between 0 and 4 points are awarded after each decision point, and the point increments reflect the ethical strength of the decision. Below is one of the situations presented in the simulation which has to do with issues of safety, responsibility, transparency, trust, and personal costs:

*Soon the data stream from the drones starts arriving and operations continue without a hitch. After a few hours, one of the most senior members of your team tells you that a rotor blade and possibly some other part of one of the drones have been broken in a crash. These accidents typically occur with sloppy landings and are something you and the team have been working on but haven't entirely been able to solve. They are probably caused by issues with usability (...) Should the local authorities be kept in the loop about any difficulties that arise with the UAVs?*

At the decision point that follows, learners need to choose one of three options: (i) not inform the local authorities and keep the technical issues to themselves, (ii) inform the local authorities once the issue has been dealt with, or (iii) inform the local authorities as soon as possible.

Another set of choices has to do with communication and the risk of excessive or false expectations. The main actor needs to decide whether to insist with local authorities that the population be clearly informed that UAVs are being used for rescue purposes. Avoiding or putting off the conversation with local authorities about informing and engaging with the community eventually results in complaints arriving to the Head of the Local Civil Protection from citizen representatives and civil society organizations; this, in turn, leads to UAV operations being halted. Discussing the issue immediately, instead, leads to another narrative segment in which the main character discusses the contents of a press release with the town hall press office. This is followed by a decision-point in which the learner is asked to choose whether and how much to emphasize in conversations with the press officers that press releases and statements to the media need to be clear about both the advantages and the limitations of using drones in search and rescue missions. Strict avoidance of hype comes at a potential personal cost, as this is the first time the main actor and their team have this type of media exposure.

## IV. USABILITY TESTING

In our study protocol, testing of the simulation is divided into three separate phases: first, usability testing; second, expert review; third, effectiveness testing. Here we describe the usability testing.

Prototype usability testing is especially important when a system is meant to be used for training, as usability issues can negatively impact on how learners interact with the system, interfering with the learning experience and ultimately hampering its effectiveness. Among methods used to assess usability that involve giving the prototype to users to interact with, two main approaches exist: observational analysis, in which users interacting with the system are observed by developers and designers, and survey-based methods, in which users are asked to fill out evaluation questionnaires after interacting with the system [26] .

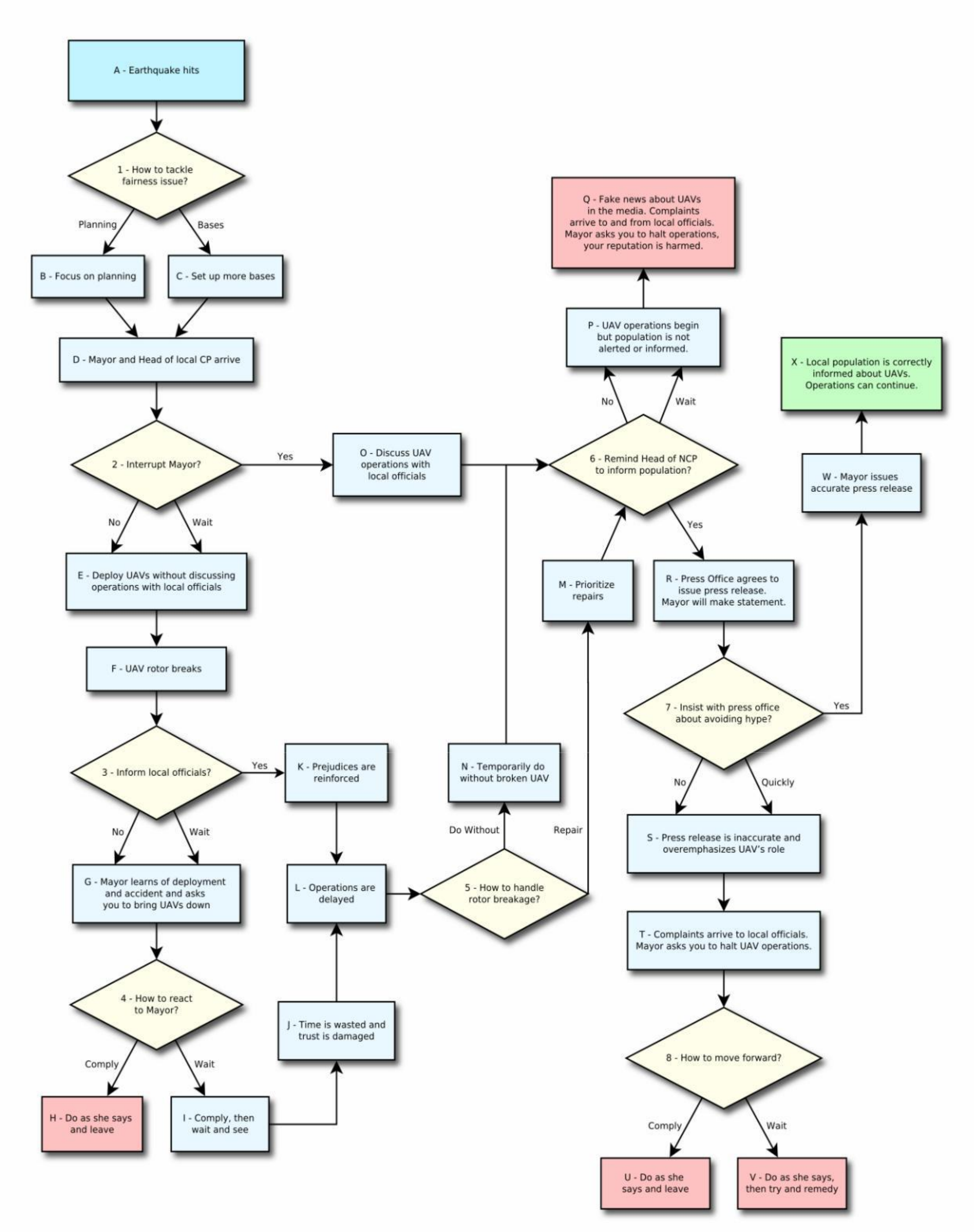


**Figure 2**. Flowchart showing the various branches and decision points of the interactive simulation.

| SUS Survey Question | Responses by Participant | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G | H | I | J | K |
| 1 I think that I would like to use this system frequently. | 4 | 4 | 3 | 5 | 4 | 5 | 5 | 4 | 5 | 4 | 4 |
| 2 I found the system unnecessarily complex. | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 I thought the system was easy to use. | 4 | 4 | 5 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 |
| 4 I think that I would need the support of a technical person to use this system. | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 5 | 1 | 1 |
| 5 I found the various functions in this system were well integrated. | 4 | 4 | 3 | 4 | 4 | 4 | 5 | 5 | 4 | 5 | 4 |
| 6 I thought there was too much inconsistency in this system. | 1 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| 7 I would imagine that most people would learn to use this system very quickly. | 4 | 4 | 5 | 5 | 4 | 5 | 4 | 5 | 5 | 4 | 4 |
| 8 I found the system very cumbersome to use. | 1 | 3 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 I felt very confident using the system. | 4 | 3 | 5 | 5 | 3 | 4 | 5 | 5 | 5 | 5 | 5 |
| 10 I needed to learn a lot of things before I could get going with this system. | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| | | | | | | | | | | | |
| **SUS Score** | **82.5** | **67.5** | **91.5** | **90** | **85** | **87.5** | **95** | **97.5** | **87.5** | **95** | **92.5** |

**Table 1**. Results of the SUS survey.

One of the most common and well-tested usability surveys is the System Usability Scale (SUS), a simple Likert-type questionnaire [27]. SUS has been extensively used with a wide range of different systems since it was first developed over twenty years ago. Moreover, it has been found to be reliable and robust, even with small numbers of users [28]. Because SUS has been used in many studies, it has been possible to establish norms that indicate a system's perceived usability [29]. While it was originally only meant to measure perceived ease of use, it has been shown to provide a global measure of system satisfaction and sub-scales of usability (items 1-3, 5-9) and learnability (items 4 and 10) [30].

Despite its advantages, SUS, like other similar metrics produces an overall usability score rather than identifying specific usability issues. When the objective is to explore how users will interact with the system and pinpoint specific problems, the most accurate approaches are observational user methods [26]. Indeed, these methods provide direct opportunities to gain insights on how users will interact with and respond to the system, in order to remediate any issues. Surprisingly, attention to remediation actions and prioritization of issues is often missing in studies [31].

Based on this background, to conduct usability testing of our text-based simulation we employed a mixed-methods approach including both a qualitative exploration and SUS. Both components of the testing were piloted with one evaluator in order to identify and eliminate any ambiguities in our approach and ensure that the pathways through the testing were clear. They were then administered to eleven evaluators (Master's and PhD students in robotics). For the qualitative element of our assessment, we used the "think aloud" method, in which evaluators were asked to verbalize their thoughts during task execution (concurrent think-aloud). Using a think-aloud protocol can help developers and designers gather evidence on how users approach tasks, understand what draws their attention, what they focus on and what they struggle with. In addition, concurrent think-aloud can capture useful, real-time information that might be lost by post-test instruments [32]. We then asked evaluators to comment on the system once they had completed execution (retrospective think-aloud). When using retrospective think-aloud, participants can focus on reporting their experience without being distracted by what they are doing [32].

Sessions with evaluators lasted between 40 and 50 minutes, totaling approximately 8.5 hours of observation; careful notes were taken throughout. Thematic analysis, a widely used method to analyze qualitative data [33], was used to identify patterns of meaning in our session notes. The results of this analysis showed that evaluators had encountered three main difficulties when trying the simulation: too much text, poor display and information overload. Evaluators reported that they were somewhat put off by the amount of text they were asked to read during the simulation. Some stated that this issue was compounded by the way the text was displayed on the screen and by the fact that the narrative segments were very information-dense, which made the story described in the simulation difficult to follow and, overall, the experience somewhat taxing from a cognitive point of view. At the same time, however, evaluators also reported that the system was simple and easy to learn, and both straightforward and pleasant to use.

For the quantitative component of our assessment, once evaluators had completed the task and retrospective think-aloud, we administered SUS. For the results of SUS, Cronbach's alpha was computed at > 0.9, suggesting excellent scale reliability, although the small sample size requires that this finding be interpreted with care. The Shapiro-Wilk test confirmed that the samples in our dataset (detailed in **Table 1**) are normally distributed. We then used the two-tailed T-test to compare our results to the literature benchmark of 80 (good user experience) [33] and found they were significantly higher (t-value: 3.336662; p = .007533, significant at p < .01.).

## V. CONCLUSIONS

We have shown here how an interactive case-based ethics simulation can be designed to support the ethics training of rescuers who may be working side by side with robots. The literature on case-based training suggests that context-specific cases like the one described in our simulation are more effective than off-the-shelf solutions in providing instructional benefits to trainees. Indeed, designing the case presented here required close collaboration between roboticists and an ethics researcher, and will require further work with other disaster robotics experts, something that is strongly advocated by proponents of Responsible Research and Innovation [34].

SUS results suggested that our system provides a very good user experience. The think-aloud protocol, however, highlighted issues with display and revealed that users perceived the story as not engrossing enough to balance out the amount of information presented. These findings provide a clear example of how crucial it is to integrate the results of usability metrics like SUS with observational user data in order to identify and correct usability problems.

The next step in our work will be to revise and re-design the simulation after carefully considering these findings. We expect to rewrite the narrative segments of the simulation to make the story more engaging by i) developing the main character and introducing other characters; ii) introducing dialogue, and iii) emphasizing drama, conflict and tension. Display issues will also be addressed, by e.g. reducing the amount of text on screen. We will then seek the opinion of disaster robotics experts and search and rescue professionals. They will be consulted about the realism of the story, of the issues presented, and of the solutions proposed in the simulation. In addition, their input will be sought about further issues of ethical relevance that could be included. Once the final version incorporating experts' and professionals' input is ready, effectiveness testing, including a case-control pre/post study and qualitative explorations, will be conducted with a sample of potential trainees.

## ACKNOWLEDGMENT

This work was supported by the Italian Ministry of Education, University and Research through the DIONISO project (Research code: B82C12000070001; National Civil Service Code SCN_00320), under the Smart Cities and Communities and Social Innovation program.

## REFERENCES

[1] R. Murphy, *Disaster Robotics*. Cambridge, Massachusetts; London, England: MIT Press, 2014.

[2] J. Jones *et al.*, "SMART Fire Fighting: The Use of Unmanned Aircraft Systems in the Fire Service," 2015.

[3] R. Murphy and S. Stover, "Rescue robots for mudslides: a descriptive study of the 2005 La Conchita mudslide response.," *J F. Robot*, vol. 25, pp. 3–16, 2007.

[4] E. Steimle, R. Murphy, M. Lindemuth, and M. Hall, "Unmanned marine vehicle use at Hurricanes Wilma and Ike," in *Unmanned marine vehicle use at Hurricanes Wilma and Ike. In: OCEANS 2009*, 2009, p. pp 1–6.

[5] R. Murphy *et al.*, "Robotics-assisted bridge inspection," *J Intell Robot Syst*, vol. 64, pp. 77–95, 2011.

[6] J. Delmerico *et al.*, "The current state and future outlook of rescue robotics," *J. F. Robot.*, vol. 36, no. 7, pp. 1171–1191, Oct. 2019.

[7] S. Geale, "The ethics of disaster management," *Disaster Prev. Manag. An Int. J.*, 2012.

[8] M. E. Gustavsson, F. K. Arnberg, N. Juth, and J. von Schreeb, "Moral Distress among Disaster Responders: What is it?," *Prehosp. Disaster Med.*, vol. 35, no. 2, pp. 212–219, 2020.

[9] M. Harbers, J. de Greeff, I. Kruijff-Korbayová, M. A. Neerincx, and K. V. Hindriks, "Exploring the Ethical Landscape of Robot-Assisted Search and Rescue," Springer, Cham, 2017, pp. 93–107.

[10] L. Battistuzzi, C. T. Recchiuto, and A. Sgorbissa, "Ethical concerns in rescue robotics: a scoping review.," *Ethics Inf. Technol.*, 2021.

[11] P. H. Werhane, "Moral Imagination and Systems Thinking," *J. Bus. Ethics*, pp. 33–42, 2002.

[12] J. L. Kolodner, *Case-based reasoning*. Morgan Kaufmann, 2014.

[13] J. L. Kolodner, J. Owensby, and M. Guzdial, "Case-based learning aids," in *Handbook of Research for Educational Communications and Technology*, 2nd ed., Mahwah, New Jersey: Lawrence Erlbaum Associated, 2004, pp. 829–861.

[14] S. Kim, W. R. Phillips, L. Pinsky, D. Brock, K. Phillips, and J. Keary, "A conceptual framework for developing teaching cases: A review and synthesis of the literature across disciplines," *Med. Educ.*, vol. 40, no. 9, pp. 867–876, 2006.

[15] J. F. Johnson *et al.*, "Case-Based Ethics Education: The Impact of Cause Complexity and Outcome Favorability on Ethicality," *J. Empir. Res. Hum. Res. Ethics*, vol. 7, no. 3, pp. 63–77, 2012.

[16] L. Falkenberg and J. Woiceshyn, "Enhancing business ethics: Using cases to reach moral reasoning," *J. Bus. Ethics*, vol. 79, no. 3, pp. 213–217, 2008.

[17] D. C. Menzel, "Teaching and Learning Ethical Reasoning with Cases," *Public Integr.*, vol. 11, no. 3, pp. 239–250, 2009.

[18] L. N. Harkrider *et al.*, "Structuring Case-Based Ethics Training: How Comparing Cases and Structured Prompts Influence Training Effectiveness," *Ethics Behav.*, vol. 23, no. 3, pp. 179–198, 2013.

[19] A. J. Plinio, J. M. Young, and L. McCormick Lavery, "The state of ethics in our society: A clear call for action," *Int. J. Discl. Gov.*, vol. 7, no. 3, pp. 172–197, 2010.

[20] T. N. Atkinson, "Using Creative Writing Techniques to Enhance

the Case Study Method in Research Integrity and Ethics Courses," *J. Acad. Ethics*, vol. 6, no. 1, pp. 33–50, 2008.

[21] Z. Bagdasarov *et al.*, "Case-Based Ethics Instruction: The Influence of Contextual and Individual Factors in Case Content on Ethical Decision-Making," *Sci. Eng. Ethics*, vol. 19, no. 3, pp. 1305–1322, 2013.

[22] C. F. Herreid, "Sorting Potatoes for Miss Bonner," *J. Coll. Sci. Teach.*, vol. 27, no. 4, pp. 236–239, 1998.

[23] C. Beal and J. Orbison, "An Interactive Professional Ethics Case Simulation," 2018.

[24] K. Vohs, R. Schmeichel, J. Twenge, N. Nelson, and D. Tice, "Making choices impairs subsequent self-control: A limited-resource account of decision making, self-regulation, and active initiative.," *J. Pers. Soc. Psychol.*, vol. 94, no. 5, pp. 883–898, 2008.

[25] L. Battistuzzi, C. Papadopoulos, T. Hill, N. Castro, B. Bruno, and A. Sgorbissa, "Socially Assistive Robots, Older Adults and Research Ethics: The Case for Case-Based Ethics Training," *Int. J. Soc. Robot.*, 2020.

[26] M. Macleod and R. Rengger, "The development of DRUM: a software tool for video-assisted usability evaluation," in *Proceedings of the 5th International Conference on Human-Computer Interaction (HCI '93),* 1993, pp. 293–309.

[27] J. Brooke, "SUS: a 'quick and dirty' usability scale.," in *Usability evaluation in industry.*, P. Jordan, B. Thomas, I. McClelland, and B. Weerdmeester, Eds. London: Taylor & Francis, 1996.

[28] A. Bangor, P. Kortum, and J. Miller, "An empirical evaluation of the System Usability Scale," *Int. J. Hum. Comput. Interact.*, vol. 24, no. 6, 2008.

[29] J. Brooke, "SUS: A Retrospective," *J. Usability Stud.*, vol. 8, no. 2, pp. 29–40, 2013.

[30] J. Lewis and J. Sauro, "Item benchmarks for the System Usability Scale," *J. Usability Stud.*, vol. 13, no. 3, pp. 158–167, 2018.

[31] R. Boring, D. Gertman, J. Joe, and J. Marble, "Proof of concept for A human reliability analysis method for heuristic usability evaluation of software," in *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society (HFES '05)*, 2005, pp. 676–680.

[32] T. Olsen, K. Procci, and C. Bowers, "Serious Games Usability Testing: How to Ensure Proper Usability, Playability and Effectiveness," in *Design, User Experience, and Usability. Theory, Methods, Tools and Practice - First International Conference, DUXU 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part II*, 2011.

[33] V. Braun and V. Clarke, "Using thematic analysis in psychology.," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006.

[34] B. Carsten and M. Coeckelbergh, "Ethics of healthcare robotics : Towards responsible research and innovation Ethics of healthcare robotics : Towards responsible research and innovation," *Rob. Auton. Syst.*, vol. 86, no. September, pp. 152–161, 2016.