Immersive Telepresence Framework
for Remote Educational Scenarios

Jean Botev[1(✉)] and Francisco J. Rodríguez Lera[2]

[1] University of Luxembourg, Av. de la Fonte 6, 4364 Esch-sur-Alzette, Luxembourg
`jean.botev@uni.lu`
[2] University of León, Campus de Vegazana S/N, 24071 León, Spain
`fjrodl@unileon.es`

Abstract. Social robots have an enormous potential for educational applications, allowing cognitive outcomes similar to those with human involvement. Enabling instructors and learners to directly control a social robot and immersively interact with their students and peers opens up new possibilities for effective lesson delivery and better participation in the classroom.

This paper proposes the use of immersive technologies to promote engagement in remote educational settings involving robots. In particular, this research introduces a telepresence framework for the location-independent operation of a social robot using a virtual reality headset and controllers. Using the QTrobot as a platform, the framework supports the direct and immersive control via different interaction modes including motion, emotion and voice output. Initial tests involving a large audience of educators and students validate the acceptability and applicability to interactive classroom scenarios.

Keywords: Social robotics · Education · Immersive telepresence · Teleoperation · Virtual reality · Human-robot interaction · UI design

1 Motivation

Social robots have an enormous potential for educational applications, allowing cognitive outcomes similar to human involvement [6]. Many research efforts focus on aspects related to autonomous and cognitive robotics for education [5,9,18,27]. However, enabling learners and instructors to directly control a social robot and immersively interact with their peers and students opens up further possibilities for effective lesson delivery, participation and tutoring in the classroom.

From an operator's point of view, which could be either a teacher or a tutor (peer) [19], the direct interaction with students is crucial to acquiring non-verbal feedback and observing immediate reactions in order to evaluate their comprehension [15]. Virtual reality (VR) technology here lends itself perfectly as an addition, complementing pure human-robot interaction (HRI) scenarios where a

robot is controlled remotely. Specifically, the combination of a VR headset with motion-based control allows the operator to more naturally translate movements to input. Together with the visual, acoustic and further channels of expression, the interaction between operator and student, despite mediated, becomes much more immersive. In addition, technologies such as emotion recognition and face detection enhance the way in which the operator perceives the student.

Educational research distinguishes various communication mechanisms between students and instructors, i.e., teachers or tutors, which include non-verbal clues that are visible to the instructor during the lesson [15,20]. These clues involve the monitoring and tracking of body movement to different extents, the time watching materials, or students looking away. Other, more subtle clues are more complex to recognize, such as blinking and the lowering of eyebrows. In both cases, many software solutions are able to capture this information from robot sensors, however, certain information and gestures are difficult to generalize and translate to software, as they are based on experiential knowledge originating from the interaction with the students.

Conversely, instructor feedback to students is equally significant to the learning process [11]. A typology based on four types of feedback extending beyond classic reward and punishment feedback by also specifying attainment and improvement is, for instance, proposed in [30]. Such feedback could be transmitted to students using verbal and non-verbal capabilities of the robot; a social robot with diverse interaction modalities thus would increase the quality and amount of feedback delivered to students.

Both of these feedback scenarios have led to several robotic telepresence solutions [8,14,35] along with a plethora of purely experimental approaches. However, this paper goes one step further in that it combines a semi-autonomous social robot with a fully immersive system based on a head-mounted display. We discuss our efforts to create an open-source framework for this kind of scenarios, as well as the exploratory experience evaluating the perception of the system before possible deployment in classrooms throughout the country of Luxembourg. To this end, we set up the fully functional platform at a national science and research fair to evaluate the interaction with students using a quick assessment method designed to capture the succinct interactions generated in events of that kind (30 pupils every 15 min in a noisy environment).

During this first iteration of our research, we intend to answer mainly the following research questions:

- Q1: What are the issues of developing an open-source software framework supported on a commercial social robot and HMD?
- Q2: How can the experience be assessed effectively using a quick evaluation method capturing a complete framework?

The remainder of this paper is structured as follows: Sect. 2 introduces the various components and technologies used in the framework, inclusive of how communication between them is handled. Section 3 then discusses important user interface (UI) and user experience (UX) design aspects along with a dedicated

assessment method, followed by the results from an initial evaluation of the system in a larger-scale context in Sect. 4. Finally, we conclude with a brief summary and outlook on future development in Sect. 5.

2   Framework and Components

This paper proposes a telepresence framework for the location-independent operation of a social robot using a VR headset and controllers as schematically illustrated in Fig. 1. Such an approach suggests a twofold communication system: on the one hand, a robot sending relevant data and receiving commands to carry out in order to interact with users and bystanders; on the other hand, a device and a dedicated user interface to present the robot's data/state to the operator and to send commands to the robot.
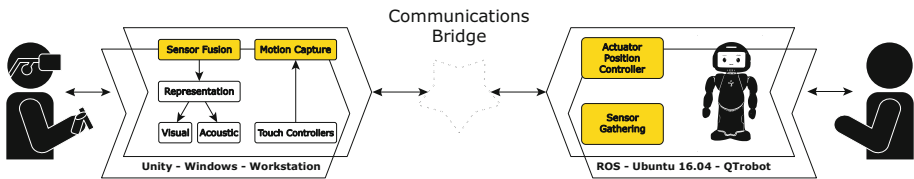

Fig. 1. Telepresence framework; VR-based operation to social robot interaction.

The telepresence framework is purposely designed to be completely location-independent, i.e., the person controlling the robot can either be in the next room or on a different continent altogether. As indicated in Sect. 1, the operator in our case is an instructor, i.e., a teacher or a tutor. Accordingly, and in line with their general use in HRI research, we employ the terms user and bystander for students and peers here [29].

In the following, we first outline the VR technology on operator side in Sect. 2.1, then introduce the robotic platform used on the user/bystander side in Sect. 2.2, and finally discuss the bridge that handles all communication between them in Sect. 2.3.

2.1   Virtual Reality

Historically, there are several definitions for VR, based either on technological means or the different notions of presence [26]. Generally speaking, VR can be described as a real or simulated environment presented to an individual through a mediating technology, so that the environment is perceived with a strong sense of presence. The feeling of immersion can be further intensified by allowing users to interact with the environment and by addressing a larger variety of sensory channels. To this day, the main channel remains visual, usually complemented by spatial audio, motion input and haptic feedback through dedicated hardware.

VR technology is composed of two central elements: the hardware, i.e., all physical components conveying the experience of and interaction with the environment, such as screens, gloves and controllers; and software allowing to develop virtual environments.

Hardware: Oculus Rift Headset and Touch Controllers. The VR hardware employed in this research are the Oculus Rift[1] headset and accompanying Touch controllers. The head-mounted display (HMD) consists of two PenTile OLED displays with an overall resolution of 2160 × 1200 at 90 Hz and a 110-degree field of view. This dual-display arrangement is complemented by two adjustable lenses which rectify the 1080 × 1200 image for each eye to create a stereoscopic 3D image. The headset features rotational and positional tracking and comes with integrated headphones supporting 3D-audio effects. The Oculus Touch controllers utilize the same low-latency tracking technology of the headset, providing a setup with joysticks and buttons for input and the opportunity for haptic feedback. Both the headset and the controllers are tracked using Oculus' Constellation sensors, a pair of external infrared cameras mounted on dedicated desk stands. The Constellation sensors, Touch controllers and Rift headset are depicted in Fig. 2.

Fig. 2. Oculus Rift VR headset, Touch controllers and Constellation sensors.

Software: Unity. Oculus provides several SDKs adding functionality to its core software. However, it can also be integrated easily with existing game engines such as Unity or the Unreal engine to harness their power for creating realistic VR experiences. Particularly Unity provides the flexibility of deploying and developing the software on a wide range of different platforms [10]. Moreover, it has a large community of developers and there are previous results where the engine has been put to good use in robot-VR scenarios [10,23]. This guarantees the scalability of the project, its long-term maintenance and platform independence.

---

[1] https://www.oculus.com/rift/.

2.2   Robotic Platform

The robotic platform utilized for this project is LuxAI's QTrobot[2], a humanoid robot with an expressive social appearance. It has a screen as its face, allowing the presentation of facial expressions and emotions using animated characters, as well as 12 degrees of freedom to present upper-body gestures. Eight degrees of freedom are motor controlled, two in each shoulder, one in each arm plus pitch and yaw movements of the head. The other four, one in each wrist and one in each hand, can be manually configured. As shown in Fig. 3 amongst other features, QTrobot has a close-range 3D camera mounted on its forehead and is provided with a six-microphone array. The QTrobot is powered by an Intel NUC processor and Ubuntu 16.04 LTS providing a native ROS interface.
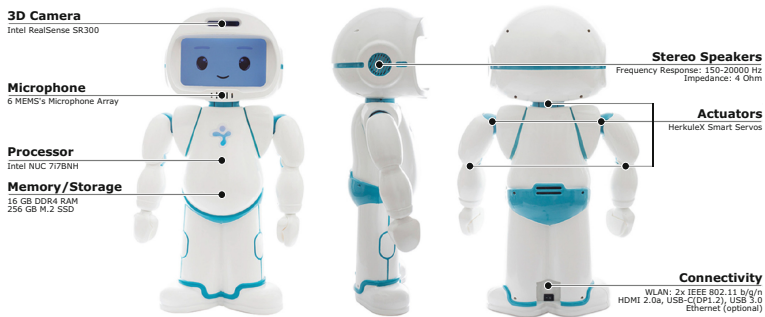

Fig. 3. QTrobot features and hardware specifications.

QTrobot behaviors can be programmed from two perspectives: a low-level programming perspective, using ROS interfaces for full robot control as in the framework implementation; and a high-level approach, through a visual programming interface presented as an Android application for tablets and smart phones, based on Blockly [21]. The latter mostly aims at users less familiar with programming, enabling them to add motion and behaviors that QTrobot can then assume when operating autonomously.

2.3   Communications Bridge

Due to the heterogeneity of VR and robot software, it is necessary to bridge the information between the different technologies. In particular, it is essential to interface the Unity engine with the ROS-based system.

As described in Sects. 2.2 and 2.1, QTrobot deploys ROS as operational middleware while the Oculus Rift uses the Unity engine; each system has its own communication protocol based on classic approaches. ROS sends messages between nodes using its own protocols called TCPROS and UDPROS. These

---

protocols generate the connections through TCP Sockets with a header containing the message data type and routing information.

For simplicity's sake, we use Rosbridge[3] for establishing communication between the QTrobot and the Oculus Rift HMD, i.e., the Unity engine. Rosbridge allows non-ROS systems to access ROS functionality by providing a JSON API. It utilizes the WebSocket protocol as a communication layer, so that any external agent with network access can send messages to the ROS environment. In particular, Rosbridge serializes ROS services and topics using a single socket interface. Figure 4 shows the bridge handling communications from the VR system to the robot and vice versa.
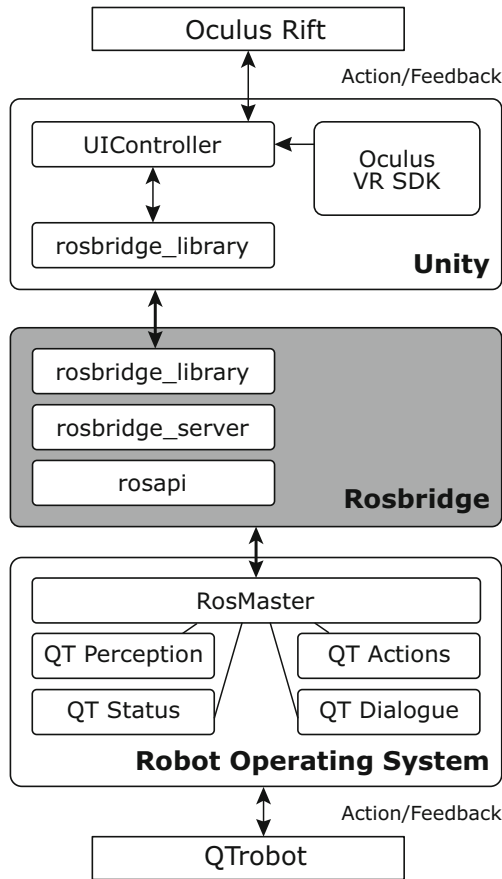

Fig. 4. Communications bridge.

Recent research related to social robotics focuses largely on artificial intelligence and autonomous robotics rather than on immersive teleoperation.

However, various approaches for interfacing VR technology with robotic systems have been proposed in different contexts, covering scenarios from fundamental HRI research aspects [33], over piloting unmanned aerial vehicles (UAV) [23], to industrial applications [25]. For example, an interesting engine for interfacing with and controlling an Arduino board is presented in [4]. It does not, however, connect to ROS. A ROS-based approach is the multi-modal man-machine communication system for remote monitoring of an industrial device as discussed in [25]. Also some UAV-related studies combine ROS with Unity [17,23], yet they often have a proof-of-concept character and focus on simulated or replicated environments rather than social HRI aspects.

The technical approach presented in this paper is an extension of the Unity-ROS interface introduced in [10]. In the same spirit, our source code is released and made available publicly via GitHub. The code has been updated to work with ROS Kinetic and Ubuntu LTS 16.04.5 Desktop which is supported until 2021. This way, it is guaranteed that other researchers can integrate our solution to their robots running ROS with only minor changes, i.e., mainly adding customized messages. Alternatively, with ROS Reality [32], a similar solution exists which aims at offering an interface for performing manipulation tasks on robots rather than predominantly social HRI control as proposed here.

The AI Robolab repository[4] provides our two packages for interfacing QTrobot with the Oculus Rift: the Unity package, called *vr_teleop*, which is the one running on the Unity SDK available in Windows; and the ROS package, called *vr_teleop_server*, which is the one interacting directly with the Unity side. In particular, it takes care of forwarding user requests to correct topics and manages the messages defining the robot status. Consequently, the repository can be considered as one of the main contributions of this study.

3   User Interface and Experience

The design of effective UIs and good UX constitutes a major area in human-computer interaction research [7]. In HRI, which is concerned with complex and dynamic control systems in real-world environments [13], the requirements tighten further as solutions for all individuals interacting with the robot or using the robot need to be considered.

UI design therefore should be seen as a process beyond the engineering development; it is a process that involves the individuals such as operators, users and bystanders, as well as the devices used during the interaction themselves. For instance, in urban search-and-rescue scenarios where different degrees of robot autonomy are supported on remote control, the UI should concentrate on four central elements [2]: 1) enhancing operator awareness of the robot environment, 2) reducing operator cognitive load, 3) minimizing the use of multiple windows/desktops, and 4) assisting with the selection of the correct robotic autonomy level in a given scenario.

---

[4] https://github.com/AIRobolab-unilu.

In a similar vein, we opted for an approach aiming at maximum immersion for the operator in order to alleviate distractions and prevent cognitive overload, as discussed in the following section.

3.1   Immersive Approach

As indicated in Table 1, there are four main elements in an HRI scenario supported on teleoperation interfaces. The various components need to be translated into the system, i.e., the scene rendering, the robotic behavior generator and the remote operation. For instance, the HRI scene is defined by the perceptive sensors integrated in the robot. The number and type of these sensors define the perception of the environment. Besides, this raw input needs manipulation and sensor fusion before it is presented to the operator. The behavior generator is associated with joint positions and the current robot status, which is relevant to a replication of the operator's movements in real time.

The goal is to create a strong sense of immersion and telepresence for the operator as defined in [12], i.e., the perception of being present in the remote environment. To achieve this, our approach suggests the transparent mapping of as many interaction channels to their direct counterparts on the operator side as possible and vice versa; e.g., visual and auditory perception (the operator sees and hears what the humanoid robot sees), or motion (the humanoid robot moves as the operator does). A VR solution composed of a headset able to relay and display robot information and a set of devices to capture operator movements (head and arms) therefore is the natural choice for a teleoperation setting as proposed here.

The HRI interface types discussed in Sects. 3.2 and 3.2 are based on previous experiences involving HMDs with controllers as well as children interacting with QTrobot, where the two modes reflect different design strategies considering individual concerns [28].

3.2   Control Modes and Requirements

When considering employing robots in remote education, all actors present in the scenario need to be taken into account: the educator, the student and any other user potentially involved in the classroom, such as teachers or tutors on site as well as student friends or colleagues. To meet the requirements of these actors, it is imperative to design a solution that involves usability guidelines for teleoperation. A validated and complete taxonomy showing a total of 70 factors associated to HRI, grouped into the following eight categories, is presented in [1]:

1. Platform Architecture and Scalability (5 factors).
2. Error Prevention and Recovery (5 factors).
3. Visual Design (10 factors).
4. Information Presentation (12 factors).
5. Robot State Awareness (10 factors).
6. Interaction Effectiveness and Efficiency (12 factors).

7. Robot Environment/Surroundings Awareness (10 factors).
8. Cognitive Factors (6 factors).

We employed the usability guidelines identified in the validated taxonomy for designing our first prototype, having in mind as actors users with little or no experience of ICT solutions. Still, the proposed control mode design explicitly considers the most frequently employed operational interfaces and the robot skills with the highest impact for the end user.

Explicit: Direct Interaction Control. In general, teleoperation solutions offer control of a platform changing between different autonomous states of the robot to allow flexibility in human control and adaptability to specific scenarios. When taking an explicit approach, the aim is to allow the operator direct control of the robot, i.e., to input and translate intention without the limitations of intermediates such as predefined options. This facilitates appropriate supervision during the interaction allowing to make informed decisions, take alternative decisions, and generate new verbal or non-verbal information, such as to any kind of motion or gestures.

Fig. 5. Explicit control mode.

The explicit control mode as implemented in our telepresence framework is shown in Fig. 5, where the operator's head and arms motion is directly forwarded and translated to the robot interacting with users and bystanders. The operator generally would be at a different location and is only in the picture for illustration purposes. Also other input, such as audio, can be forwarded and processed

in real time. Without such direct interaction control and the use of different sensory channels, achieving immersion and telepresence as defined in [12] would be impossible, or at least much harder, to achieve.

Implicit: Social Interaction Control. In addition to the explicit control mode, in an educational context it is imperative to offer a complementary type of control which focuses on human-robot and human-human roles as well as their relationships. This way, the robot can offer predefined voices and prerecorded sentences, gestures or expressions provided not only by educational stakeholders, such as teachers and tutors, but also by individuals with a personal connection, such as family members or friends. Depending on the specific situational context, the operator selects from a set of associated, implicit options.
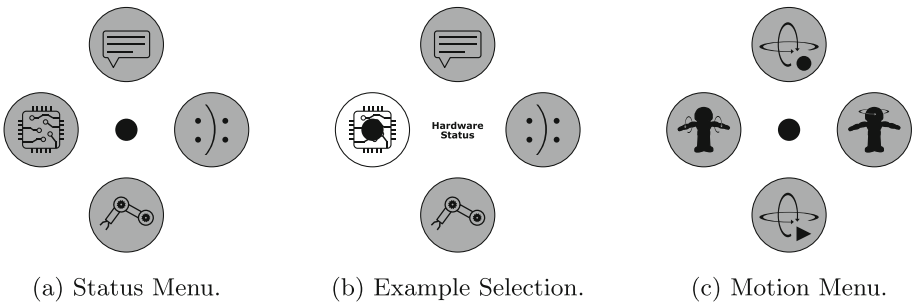

(a) Status Menu.   (b) Example Selection.   (c) Motion Menu.

Fig. 6. Implicit control mode - UI overlays for access to available functions.

The different UI overlays defined in our framework are shown in Fig. 6. When triggered, the different options are displayed in a circle and the operator can use the analog stick to navigate to and select a specific item. In Fig. 6b, for instance, the operator is about to select the hardware option from the status menu to retrieve relevant system information. The status menu is triggered from the left Touch controller. The other options in the status menu include motor/servo status (bottom), as well as dialog (top) and robot motivational variables (right) component controls. The motion menu in Fig. 6c is triggered from the right Touch controller. It provides access to the various motion-related options, e.g., to record (top) or play (bottom) complex motion sequences, as well as fine-grained control of QTrobot's arms (left) or head (right). The latter are displaying options related to the conversation and facial expression modules.

The status information presented to the operator wearing the Oculus HMD has a variety of sources: motor/servo status reflects the different motor positions within a normalized numeric range. The dialog menu opens a new window where a set of variables and facial expressions can be selected and triggered using the QTrobot speech system and display. Motivational variables indicated in the corresponding menu constitute a subset of the variables described in [22] to outline the current mood of the robot: curiosity, frustration, fatigue and pain. This particular option is shown in the screenshot in Fig. 7.
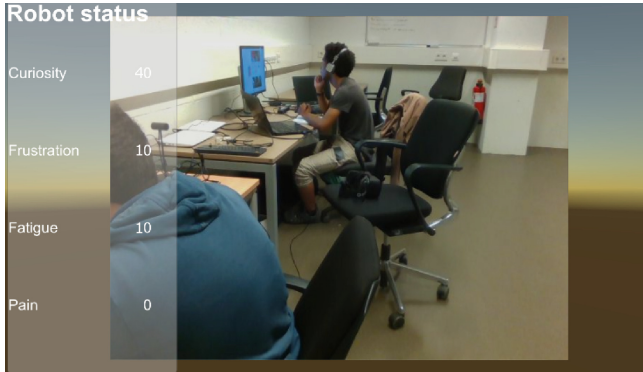
Fig. 7. Transparent status information overlay on reduced camera viewport.

3.3   Measuring UI/UX Impact

Beyond the lab, any kind of HRI scenario suffers from variability and dynamic changes in the environment. When developing the UI design, a number of factors defining HRI scenarios need to be considered because they influence or directly affect interaction [34]:

1) Ratio of People to Robots,
2) Amount of Intervention,
3) Level of Shared Interaction Among Teams,
4) Composition of Robot Team,
5) Human-Robot Physical Proximity,
6) Decision Support for Operators,
7) Space,
8) Interaction Roles,
9) Criticality,
10) Time,
11) Autonomy,
12) Task Type, and
13) Robot Morphology.

These categories comprise most part of the HRI elements for the different actors defining the HRI scenario [3]:

– HRI system: the technological elements involved during interaction, mainly composed of the following three elements: the system that perceives information from and presents information to individuals, which is a robot; the communication system, focusing on the network transmission; and the display system, which makes the operator perceive the remote scene and generate actions to influence bystanders.
– Robot: the mobile or non-mobile agent that interacts in one way or another with the bystander.

– Remote operator: the individual having the capacity to change or manipulate the robot behaviors so as to provide the correct actions in a given scenario. The set of actions is proposed by a supervisor.
– Supervisor: the individual who monitors and controls the overall situation during the HRI scenario, also known as commander or director.
– Bystander: the individual or individuals in the remote environment whose behaviors affect the robot actions (both autonomous and via remote operator).

A good design should enhance the effectiveness, efficiency and efficacy of an HRI approach, but the evaluation is not a straightforward process. In educational environments, this assessment is further complicated by potential ethical and technological implications of working with actual students and professional educators. Finding issues in the system only during the experimental phase might turn costly for both the researchers and the bystanders involved in the process.

Based on validated approaches from the mapping review of the literature that we performed, we propose a set of parameters to be evaluated for measuring UI/UX impact as in Table 1, defining the Quick Assessment Method (QAM). Each parameter has an effect on particular actors and is therefore evaluated independently. The proposed parameters have well-defined metrics, however, interviews with the director and developer can further accelerate evaluation and development. At this point, these interviews assess the impact following a Likert scale that focuses on the level of influence: 1, not at all influential; 2, slightly influential; 3, somewhat influential; 4, very influential; 5, extremely influential.

Table 1 contains four elements: the parameter ID, the parameter name, the expected effect, supported on the proposed scale, and the perceived effect, using the same scale, which is obtained once the design is evolving or tested during development and pre-experimental phases. The expected and perceived parameter values in the table indicate the values obtained during a first evaluation of the proposed telepresence framework discussed in Sect. 4.

4   Evaluation

An initial evaluation of the system in action was conducted at the recent FNR Researchers' Days[5], a two-day national science and research fair. The sixth edition in November/December 2018 attracted more than 6,500 attendees. Visitors from the mostly laymen audience continuously tried the system at our stand.

The photo in Fig. 8 shows part of the stand. Visitors were not only allowed to interact with the robot, but also to take on the role of its operator. The setup included a projector displaying the image routed to the HMD together with other data to the remaining audience, so as to give a detailed impression of what is being captured by the robot's sensors and actuators.

---

[5] http://www.researchersdays.lu/.

Table 1. Quick Assessment Method (QAM), cf. [3, 31].

ID	Parameter	Exp.	Perc.
HRI System			
1	Percentage of requests for assistance made by robot	1	1
2	Percentage of requests for assistance made by operator	4	5
3	Number of interruptions of operator rated as non-critical	3	4
4	Effectiveness (objective ratings)	-	-
5	Efficiency (objective ratings)	-	-
6	Efficacy (objective ratings)	-	-
7	Ease of use (subjective rating)	5	5
8	Ease of learning (subjective rating)	5	5
9	Functional primitives decomposition	5	5
10	Interaction effort	3	2
Robot			
11	Robot self-awareness	1	3
12	Human awareness	1	1
13	Autonomy	1	3
Remote Operator			
14	Human-robot situation awareness	5	5
15	Human-human situation awareness	2	1
16	Robot-human situation awareness	5	5
17	Robot-robot situation awareness	1	1
18	Human's overall mission awareness	3	5
19	Robot's overall mission awareness	1	1
20	Workload	3	3
21	Accuracy of mental models of device operation	5	5
Bystander (Student, Educational Professional)			
22	Human-robot situation awareness	1	1
23	Human-human situation awareness	1	1
24	Accuracy of mental models for interaction	3	5
25	Workload	1	1
26	Social acceptance	3	5
27	Effort expectancy	5	5
28	Attitude toward using technology	5	5
29	Self-efficacy	3	5
30	Attachment	1	1
31	Reciprocity	1	1
32	Emotion	3	5
33	Feeling of security	5	5

Fig. 8. System demonstration and assessment at national science and research fair.

Sociodemographic Observations. During the first day, attendance of the fair is limited to high-school pupils, journalists and invited guests, while the second day is open to the general public. 120 school classes registered to the event with, according to the organization recap, a total of 1,907 pupils (excluding teaching staff) visiting the event over the first day. The visits were peaking between 9:30 and 12:00. The Luxembourgish pupils were between 12 and 19 years old. Our stand was visited by approximately one group of students per 30 min, however, every 5–10 min a student or teaching personnel was operating the robot for 5-min time slots during which they interacted with friends and other visitors. The presentations were held in different languages, with the main languages used being German and French, corresponding with the main tongues spoken by visitors. The main language available in QTrobot is English. The second day of the fair comprised around 4,500 visitors with no age limit. The team proceeded with the same established presentation mode of the first day. At the end of each day, there were around 50–60 system tests with pupils, children, parents and teachers contributing to the overall evaluation data.

Objective Observations. Despite operator changes in quick succession and constant operation, the robot was running robustly most of the time. The VR system was used by one visitor every five to ten minutes throughout the day, i.e., around ten individual operators per hour, eight hours per day. Due to the extremely busy environment, there were network issues when using Wi-Fi connections; using Ethernet greatly improved delays, however, some audio delays are related to the first version of the communications bridge, since it delivers the signal in chunks. In terms of UX, due to the minimal interface, there was close to no setting-in period even for users yet unfamiliar with VR or controller-based input. The interaction was natural, and the degree of immersion high,

with users putting themselves readily in QTrobot's position to interact with the person opposite.

Subjective Observations. The feedback on both operation and reception sides was very positive and encouraging. There was a strong feeling of presence [24] and embodiment [16], with only few operators displaying slight symptoms of disorientation. The robot was accepted without exception by young and old, with highly focused interaction in particular also from younger children. This group was particularly unbiased towards the robot, however, the visitors generally interacted with the robot as if it were the most natural thing in the world.
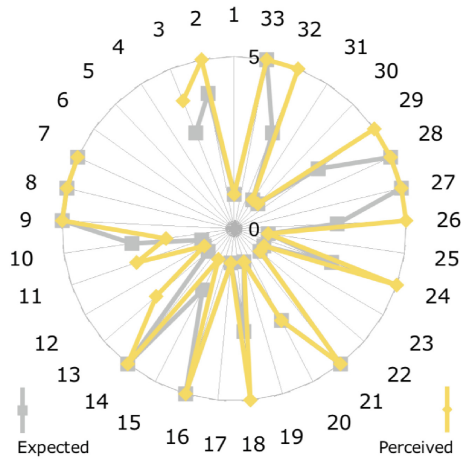

Fig. 9. Expected vs perceived results using the QAM.

QAM Observations. Following the evaluation of the objective and subjective feedback obtained during the tests at the Researchers' Days, the data was correlated with the QAM parameters to show the expected versus the perceived results. The visualization in Fig. 9 clearly indicates that the final users met our expectations and indeed were comfortable with the immersive approach. In particular, parameter 14 (human-robot situation awareness) confirms our approach for reducing the interaction effort, with the operators perceiving low complexity in direct interaction control mode. While deliberately employing an expressive humanoid robot to reduce bystander's reservations about the technology, the high social acceptance (parameter 26) and self-efficacy (parameter 29) assessment results greatly exceeded our expectations. Other factors, such as those related to particular robot skills, indicate that it is important to show also recurring, autonomous behaviors in the robot, as bystanders were expecting familiar or movie-like actions.

5   Conclusion and Outlook

This paper introduces a telepresence framework for the location-independent operation of a social robot using a VR headset and controllers with motion-tracking technology. Such robots have an enormous potential for educational applications: different to employing autonomous techniques, enabling learners and instructors to directly and immersively control a platform like QTrobot and interact with users and bystanders opens up new possibilities for effective participation and tutoring. Throughout, the robot operators exhibited a strong feeling of presence and embodiment.

Besides detailing the technical implementation of the framework along with the various interaction modes, this paper also discusses a multi-phase evaluation approach to assess HRI scenarios in different contexts, including an initial evaluation of the proposed framework. The preliminary tests involved a large and heterogeneous audience, and already helped validating the general applicability to interactive educational settings.

Our vision includes the full experimental evaluation of our approach in ordinary classrooms with professional educators on operator side and regular students on bystander side in order to push forward the inclusion of interactive, social robots and remote presence technologies in schools. Before doing so, however, we would like to further improve the network robustness and optimize servo/motor control, potentially with additional gesture input. Another possible extension would be to add a UI elements to mitigate the disorientation issue we observed at times during the demonstration of the system.

References

1. Adamides, G., Christou, G., Katsanos, C., Xenos, M., Hadzilacos, T.: Usability guidelines for the design of robot teleoperation: a taxonomy. IEEE Trans. Hum.-Mach. Syst. **45**(2), 256–262 (2015)
2. Baker, M., Casey, R., Keyes, B., Yanco, H.A.: Improved interfaces for human-robot interaction in urban search and rescue. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC 2004), pp. 2960–2965 (2004)
3. de Barros, P.G., Linderman, R.W.: A survey of user interfaces for robot tele-operation. Technical report, Worcester Polytechnic Institute (2009). http://digitalcommons.wpi.edu/computerscience-pubs/21
4. Bartneck, C., Soucy, M., Fleuret, K., Sandoval, E.B.: The robot engine - making the unity 3D game engine work for HRI. In: Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015), pp. 431–437 (2015)
5. Belpaeme, T., et al.: Child-robot interaction: perspectives and challenges. In: Herrmann, G., Pearson, M.J., Lenz, A., Bremner, P., Spiers, A., Leonards, U. (eds.) ICSR 2013. LNCS (LNAI), vol. 8239, pp. 452–459. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-02675-6_45

6. Belpaeme, T., Ramachandran, A., Scassellati, B., Tanaka, F.: Social robots for education: a review. Sci. Robot. **3**(21) (2018)
7. Benyon, D.: Designing Interactive Systems: A Comprehensive Guide to HCI, UX and Interaction Design. Pearson Edinburgh (2014)
8. Cha, E., Chen, S., Matarić, M.J.: Designing telepresence robots for K-12 education. In: 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2017), pp. 683–688 (2017)
9. Clabaugh, C., Matarić, M.: Escaping Oz: autonomy in socially assistive robotics. Ann. Rev. Control Robot. Auton. Syst. **2**, 33–61 (2019)
10. Codd-Downey, R., Forooshani, P.M., Speers, A., Wang, H., Jenkin, M.R.M.: From ROS to unity: leveraging robot and virtual environment middleware for immersive teleoperation. In: Proceedings of the 11th IEEE International Conference on Information and Automation (ICIA 2014), pp. 932–936 (2014)
11. Crooks, T.J.: The impact of classroom evaluation practices on students. Rev. Educ. Res. **58**(4), 438–481 (1988)
12. Draper, J.V., Kaber, D.B., Usher, J.M.: Telepresence. Hum. Factors **40**(3), 354–375 (1998)
13. Fong, T., Thorpe, C., Baur, C.: Collaboration, dialogue, human-robot interaction. In: Jarvis, R.A., Zelinsky, A. (eds.) Proceedings of the 10th International Symposium on Robotics Research (ISRR 2003), pp. 255–266 (2003)
14. Gallon, L., Abenia, A., Dubergey, F., Négui, M.: Using a telepresence robot in an educational context. In: Proceedings of the 10th International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS 2019), pp. 16–22 (2019)
15. Jecker, J.D., Maccoby, N., Breitrose, H.: Improving accuracy in interpreting nonverbal cues of comprehension. Psychol. Sch. **2**(3), 239–244 (1965)
16. Kilteni, K., Groten, R., Slater, M.: The sense of embodiment in virtual reality. Presence: Teleoperators Virtual Environ. **21**(4), 373–387 (2012)
17. Meng, W., Hu, Y., Lin, J., Lin, F., Teo, R.: ROS+Unity: an efficient high-fidelity 3D multi-UAV navigation and control simulator in GPS-denied environments. In: 41st Annual Conference of the IEEE Industrial Electronics Society (IECON 2015), pp. 2562–2567 (2015)
18. Miller, D.P., Nourbakhsh, I.: Robotics for education. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, pp. 2115–2134. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32552-1_79
19. Mubin, O., Stevens, C.J., Shahid, S., Al Mahmud, A., Dong, J.: A review of the applicability of robots in education. Technol. Educ. Learn. **1**, 1–7 (2013)
20. Okon, J.: Role of non-verbal communication in education. Mediterranean J. Soc. Sci. **2**(5), 35–40 (2011)
21. Pasternak, E., Fenichel, R., Marshall, A.N.: Tips for creating a block language with blockly. In: Proceedings of the IEEE Blocks and Beyond Workshop (B&B 2017), pp. 21–24 (2017)
22. Rodríguez-Lera, F.J., Matellán-Olivera, V., Conde-González, M.Á., Martín-Rico, F.: HiMoP: a three-component architecture to create more human-acceptable social-assistive robots. Cogn. Process. **19**(2), 233–244 (2018)
23. Roldán, J.J., Peña-Tapia, E., Garzón-Ramos, D., de León, J., Garzón, M., del Cerro, J., Barrientos, A.: Multi-robot systems, virtual reality and ROS: developing a new generation of operator interfaces. In: Koubaa, A. (ed.) Robot Operating System (ROS). SCI, vol. 778, pp. 29–64. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-91590-6_2

24. Schuemie, M.J., van der Straaten, P., Krijn, M., van der Mast, C.A.: Research on presence in virtual reality: a survey. CyberPsychol. Behav. **4**(2), 183–201 (2001)
25. Sita, E., Horváth, C.M., Thomessen, T., Korondi, P., Pipe, A.G.: ROS-Unity3D based system for monitoring of an industrial robotic process. In: Proceedings of the 10th IEEE/SICE International Symposium on System Integration (SII 2017), pp. 1047–1052 (2017)
26. Steuer, J.: Defining virtual reality: dimensions determining telepresence. J. Commun. **42**(4), 73–93 (1992)
27. Toh, L.P.E., Causo, A., Tzuo, P.W., Chen, I.M., Yeo, S.H.: A review on the use of robots in education and young children. J. Educ. Technol. Soc. **19**(2), 148–163 (2016)
28. Tromp, N., Hekkert, P., Verbeek, P.P.: Design for socially responsible behavior: a classification of influence based on intended user experience. Des. Issues **27**(3), 3–19 (2011)
29. Tsui, K.M., Desai, M., Yanco, H.A.: Considering the bystander's perspective for indirect human-robot interaction. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010), pp. 129–130 (2010)
30. Tunstall, P., Gipps, C.: Teacher feedback to young children in formative assessment: a typology. Br. Educ. Res. J. **22**(4), 389–404 (1996)
31. Weiss, A., Bernhaupt, R., Lankes, M., Tscheligi, M.: The USUS evaluation framework for human-robot interaction. In: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction at the Adaptive and Emergent Behaviour and Complex Systems Convention (AISB 2009), pp. 11–26 (2009)
32. Whitney, D., Rosen, E., Phillips, E., Konidaris, G., Tellex, S.: Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In: Amato, N.M., Hager, G., Thomas, S., Torres-Torriti, M. (eds.) Robotics Research. SPAR, vol. 10, pp. 335–350. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-28619-4_28
33. Whitney, J.P., Chen, T., Mars, J., Hodgins, J.K.: A hybrid hydrostatic transmission and human-safe haptic telepresence robot. In: 22nd IEEE International Conference on Robotics and Automation (ICRA 2016), pp. 690–695 (2016)
34. Yanco, H.A., Drury, J.: Classifying human-robot interaction: an updated taxonomy. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (ICSMC 2004), vol. 3, pp. 2841–2846 (2004)
35. Zhang, M., Duan, P., Zhang, Z., Esche, S.: Development of telepresence teaching robots with social capabilities. In: Proceedings of the ASME International Mechanical Engineering Congress and Exposition (IMECE 2018), pp. 1–11 (2018)