Advanced Robotics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tadr20

Human-robot interactions with an autonomous health screening robot in long-term care settings

Cristina Getson & Goldie Nejat

To cite this article: Cristina Getson & Goldie Nejat (2023) Human-robot interactions with an autonomous health screening robot in long-term care settings, Advanced Robotics, 37:24, 1576-1590, DOI: 10.1080/01691864.2023.2293133

To link to this article: https://doi.org/10.1080/01691864.2023.2293133

Published online: 18 Dec 2023.

ADVANCED ROBOTICS
2023, VOL. 37, NO. 24, 1576–1590
https://doi.org/10.1080/01691864.2023.2293133

FULL PAPER

Human-robot interactions with an autonomous health screening robot in long-term care settings

Cristina Getson and Goldie Nejat

Autonomous Systems and Biomechatronics Lab, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

ABSTRACT
Socially assistive robots are increasingly being considered to help address the shortage of care workers in long-term care, which has been further exacerbated by the COVID-19 pandemic. In this paper, we present the first human-robot interaction study with care staff and an autonomous screening socially assistive robot in a long-term care facility. We assessed: (1) overall perceptions, experiences and attitudes of care staff prior to and after interacting with the robot, and (2) perceived workload and usability of the robot by administrators and management staff. Results show staff had overall high ratings of the robot, with a statistically significant increase identified for cognitive attitude towards the robot after interaction. Furthermore, we found that overall, perceived workload was moderately low as defined by the NASA Task Load Index while using the robot screener, and the usability rating of the robot was rated between *OK* and *Good* by the System Usability Scale. Personalization of the robot was found to be an important factor for usability. Staff enjoyed using the robot and had high willingness to frequently use it. In general, our robot study motivates the application of autonomous socially assistive robots from the staff perspective for repetitive tasks in long-term care homes.

ARTICLE HISTORY
Received 22 December 2022
Revised 12 September 2023
Accepted 27 November 2023

KEYWORDS
Human-robot interaction; assistive robotics; care workers; health screening; workload and usability in robot use

1. Introduction

The need for care workers has been exceeding the human resources available [1,2], with increasing demands placed on existing healthcare staff in long-term care (LTC) due to a growing older population [3]. Furthermore, staff shortages and high turnover rates have further escalated this need during the COVID-19 pandemic [4,5]. Technological solutions, including robots, are needed for LTC to meet the increasing demands on staff [1].

During the pandemic, socially assistive robots (SARs) were used to help minimize human-human contact, to ensure risk-free environments, and to promote physical and psychological well-being [6,7]. Additionally, they have helped alleviate staff workload and ensured staff safety by performing functions such as receptionist, telepresence to communicate between residents and family members or medical staff, and monitoring the health of residents [8]. Our own prior research included the development and deployment of an autonomous SAR for COVID-19 screening of staff working in a LTC home, where we introduced the robot screener to support the human screening process, and investigated staff's overall interaction experiences with, attitudes toward, and acceptance of a socially assistive robot for the screening task [9]. This prior study investigated staff perceptions and acceptance of a social robot over an extended period of time, and within the context of technology adoption. Our objective was not for the robot to replace a human, but rather to provide an additional resource to help with the screening task during the busy shift commencement times.

In this paper, we extend our research work on our screening robot by further investigating and analyzing the perceived workload and usability of utilizing a SAR for the pertinent screening task to ensure uptake of the technology during and post-COVID. Both workload and usability with respect to the introduction and direct use of SARs by staff in LTC homes for robot-staff interactions during the pandemic have not yet been explored; these are important factors in assessing ease of use and the willingness of staff to adopt and use SARs. When a new technology is first introduced, such factors can act as barriers to uptake [10]. This research contributes to the design and implementation of SARs as an additional resource in a long-term care setting from the staff perspective, by investigating the ease of use, usability, and impact on staff workload during a high-stress time when increasing work demand was placed on staff members.

CONTACT Cristina Getson cristina.getson@mail.utoronto.ca

Additionally, it contributes to the emerging but growing body of research exploring social robot-assisted human-interaction tasks in healthcare settings, where staff need to complete multiple tasks.

Staff technology readiness, defined as 'preparedness and willingness to accept and use new technology to achieve goals at work' [11], is a key consideration in human-robot interactions (HRI) [12]. Furthermore, organizational support is a key element in promoting technology readiness [13]. The mediation of the incorporation and use of robotic technology by main managerial and administrative stakeholders at care facilities helps reconcile barriers to implementation, such as technical issues and complexity of use [14].

2. Related work

The use of robots by staff in long-term care settings pre-pandemic has mainly focused on attitudes and acceptance of robots providing entertainment such as Bingo [15], practical assistance [16], and health monitoring [17]. Herein, we categorize and discuss HRI studies that have considered the workload and/or usability of SARs by care workers.

2.1. Socially assistive robot workload studies with care workers

Workload can be defined as the cost incurred by humans when performing a task in order to achieve a specific level of performance [18]. It is influenced by the task requirements, the circumstances under which it is performed, and the skills, behaviors, and perceptions of the humans involved [18]. In social HRI, workload has mainly been measured using either semi-structured interviews and observations, such as in [19,20], and questionnaires, such as in [15]. We discuss workload as evaluated by care workers when using SARs.

In [19], care worker usage of the Pepper robot was observed in a care home during the pandemic. A range of robot psycho-social activation (age guessing, playing songs) and cognitive activation (quiz games) applications were made available for care workers to choose from for robot interactions with residents. No specific instructions on how to use the robot were provided. Observations of care workers and management staff, along with semi-structured interviews, and log files of robot usage time and usage patterns were collected. Care workers mainly used the robot by placing it in residents' rooms for one-on-one interactions. Interviews with staff showed that the robot did not reduce staff workload in terms of personnel or time resources, however, it was perceived as a useful tool to assist care workers in their daily work with residents.

In [20], the views and attitudes of care workers on the use of robots in care homes were obtained through semi-structured interviews. The care workers had previously observed Pepper interacting with residents by entertaining them with music, videos, playing games, telling jokes, displaying the news, or providing reminders, as outlined in [21]. In general, staff had positive views about the robot and saw robots as having the potential to be supplementary tools to human carers, particularly by sharing workload and helping to improve care already provided.

In [15], care workers, without any robotic experience, taught the socially assistive robot Tangy how to facilitate recreational activities such as Bingo games through learning from demonstration. A 3-part post-interaction questionnaire was administered to care workers, measuring user experience through open-ended questions, perceived workload through the NASA Task Load Index (NASA-TLX) questionnaire [18], and perceived usability through the System Usability Scale (SUS) questionnaire [22]. The results showed moderately low perceived workload for teaching and personalizing the robot's verbal and nonverbal facilitation behaviors.

2.2. Socially assistive robot usability studies with care workers

The Standards Organization (ISO 9241-11) [23] defines usability as 'the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.' In HRI, usability takes into account the overall quality of a user's experience when interacting with a SAR [24]. It deals with user perception of the interface, the interaction, and the outcome [25]. Robot usability in HRI scenarios with care workers has been evaluated through interviews and questionnaires, such as in [26,27].

In [26], the usability of the small-size robot NAO was compared to that of a tablet in helping older adults in an elderly care facility complete a set of health monitoring and physical training tasks. For each technology, the older adults completed the SUS along with 5-point Likert scales for perceived usefulness, enjoyment, and control. They then stated their preference (for NAO or the tablet) via a short interview. The SUS scores, (representing a *D* on the grade scale as interpreted in [28]), showed that both technologies suffered from usability issues, such as not hearing the robot and responding to it in time. The scores for perceived usefulness and enjoyment were positive for both technologies. However, a third of the participants preferred the NAO robot over the tablet for monitoring and health training.

In [27], a customized software version of the NAO robot, called Zora, was used by care workers to stimulate physical activities of older adults in LTC. Care workers evaluated the robot's usability through a modified Usability, Satisfaction, Ease of Use (USE) questionnaire. Qualitative data were also obtained through observations and interviews. The results showed that 67% of care workers experienced more fun at work when using Zora. The majority of them indicated they were happy when working with Zora and believed that the residents were content when using Zora.

In [15], the SUS questionnaire was used to evaluate perceived usability of care workers in teaching the Tangy robot how to facilitate Bingo games. The mean SUS scores were interpreted with an *OK* rating. Improvements were suggested with respect to comfort by providing more teaching trials on the system and increasing the teaching speed for the activity.

2.3. Summary

A handful of studies have quantitatively measured both workload and usability in social HRI. In [29], mental workload and usability were measured for a teleoperation task, however, the remote users were not care workers. With respect to caregivers, in [15] workload and usability were measured for using a learning by demonstration system to teach the robot to assist residents, rather than caregivers' direct interactions with the robot.

Whereas our application focuses on the use of an autonomous SAR for direct staff-robot social interactions to help staff with a required task administered by the robot (health screening). To the authors' knowledge, no studies have yet been conducted that investigate workload and usability of a SAR by care staff in LTC during the pandemic using standardized measurement scales, such as NASA-TLX and SUS. Due to increasing interest in deploying SARs to help with staff shortages in LTC settings [30], there is a need to investigate how these robots impact workload and how their usability affects caregivers.

3. Research questions

This work addresses the following research questions:

- RQ1: Does interacting with a SAR for screening have minimal impact on staff workload, as measured by the NASA-TLX?
- RQ2: Does robot perceived usability, as measured by the SUS, impact the willingness of staff to use the SAR?

4. Autonomous socially assistive robot COVID-19 screening study

Our objective was to conduct an exploratory HRI study during the COVID-19 pandemic to investigate the utilization and effectiveness of a social interactive screening robot in a LTC setting. This study took place in the Fall of 2021. We evaluated staff members' experience with the robot, as well as how demographics influence staff attitudes. The robot screening study took place at a LTC home in Toronto, Canada, with the Pepper robot over the course of two months. The study was approved by the University of Toronto's Ethics Board.

4.1. Participants

From approximately 200 staff members at the LTC home, 84 participants signed up for the study. These included administrators, nurses, personal support workers, and those working in rehabilitation and social care. Each participating staff member was given a unique QR code to use during the robot screening task to maintain anonymity for the purpose of the study. Staff were recruited through: (1) posting flyers throughout the home with information about the study, (2) emails sent from the home administrators introducing the study, and (3) a demo presentation of the screener robot at the home.

4.2. Robot design

Participants interacted socially with the Pepper robot in an autonomous way. A contactless thermometer was placed next to the robot, along with a box of masks on a table, as seen in Figure 1. The robot used the QR code to anonymously keep track of all the people screened. The robot asked screening questions provided by the Ontario Ministry of Health and responses were recorded in a CSV file, which was automatically emailed to an administrative staff member at the LTC home.

A graphical user interface (GUI) was developed using HTML for Pepper's tablet, which displayed corresponding text for the robot's speech, images of the person with/without a mask, an image of the detected QR code, and a progress bar at the top of the screen to indicate screening progress. A video of the screening procedure with Pepper can be found on our Youtube channel: https://www.youtube.com/watch?v=X6EKXENu9bY.

Screening events are recorded in a CSV file, as presented in Figure 2, which includes a time-stamp, temperature confirmation, face mask confirmation, unique QR code, and answers to the four health screening questions (found in the Appendix). This CSV file is automatically emailed by the robot to administrative staff at the facility.

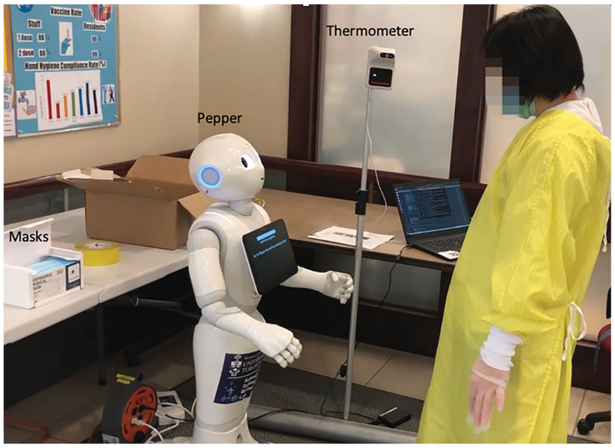


Figure 1. Set-up at the front entrance of the long-term care home, with the Pepper robot, masks, contactless thermometer, and a staff member.

A Lenovo Thinkpad P1 Gen3 laptop running Ubuntu 18 was used to control the robot using software we developed in Python. We used an off-the-shelf mask detection software, AIZOOTech FaceMask Detection algorithm, which uses Convolutional Neural Networks, available on Github [31]. The algorithm, composed of parts of WIDER Face [32] and Masked Faces (MAFA) datasets [33], is able to classify faces with masks (with an average accuracy of 91.9%) and without masks (average accuracy of 89.6%) [31]. Within the class 'faces with mask', there are two sub-classes: (1) correctly worn, and (2) incorrectly worn, where the former detects if the mask covers both the nose and the mouth. This detection algorithm has been used by Softbank Robotics for the Pepper robot running QiSDK [34], and by several researchers [35,36].

4.3. Procedure

The robot was placed at the front entrance of the LTC home, to screen care workers as they arrived to start their shifts, at both 6:30 am and 2:30 pm. Management and admin staff arrived at 9:00 am to undergo screening. The robot interacted and asked questions autonomously; screening steps are presented in Figure 3 and Table 1. If the robot screening failed due to any of the following conditions: (1) temperature $>$ 37.5°C, (2) a health screening question was answered with 'Yes', or (3) a missing or non-registered QR code; then the robot sounds an audible alarm and Pepper would ask them to go see a staff member at reception, also situated at the entrance.

The mask detection method detects if multiple people are within the field of view of Pepper's RGB camera; if this is true, the robot will instruct them to keep their social distance. Then, mask detection will take place only after the robot requests that the person takes a mask from the box of masks provided on the adjacent desk, to ensure a smooth flow of the screening process. A visual confirmation is used to let the person know if their mask is properly worn with Pepper displaying an image of the person's face on its screen, to further provide instructions. This image of the person is only shown on the screen and is not stored by the robot to ensure anonymity of the participants.

4.4. Measures

For all staff: Pre- and post-study 5-point Likert questionnaires were completed by participants, before interacting with the screening robot, and at the end of the study, after having interacted at least once with the robot. These

1	time	temperature	mask_on	qrcode	Q1	Q2	Q3	Q4
2	21-12-13-06:29	No Fever	Mask Detecte	Employee #1	no	no	no	no
3	21-12-13-06:47	No Fever	Mask Detecte	Employee #44	no	no	no	no
4	21-12-13-07:02	No Fever	Mask Detecte	Employee #3	no	no	no	no
5	21-12-13-08:21	No Fever	Mask Detecte	Employee #1	no	no	no	no
6	21-12-13-08:26	No Fever	Mask Detecte	Employee #17	no	no	no	no
7	21-12-13-08:52	No Fever	Mask Detecte	Employee #25	no	no	no	no
8	21-12-13-08:54	No Fever	Mask Detecte	Employee #7	no	no	no	no
9	21-12-13-09:13	No Fever	Mask Detecte	Employee #6	no	no	no	no
10	21-12-13-09:16	No Fever	Mask Detecte	Employee #4	no	no	no	no
11	21-12-13-09:26	No Fever	Mask Detecte	Employee #75	no	no	no	no
12	21-12-13-09:59	No Fever	Mask Detecte	Employee #83	no			
13	21-12-13-10:01	No Fever	Mask Detecte	Employee #83	no	no	no	no
14	21-12-13-14:32	No Fever	Mask Detecte	Employee #84	no	no	no	no
15	21-12-13-14:34	No Fever	Mask Detecte	Employee #60	yes	no	no	no

Figure 2. CSV file sample with screening time stamp, temperature, mask, QR code confirmation, and answers to the 4 screening questions.

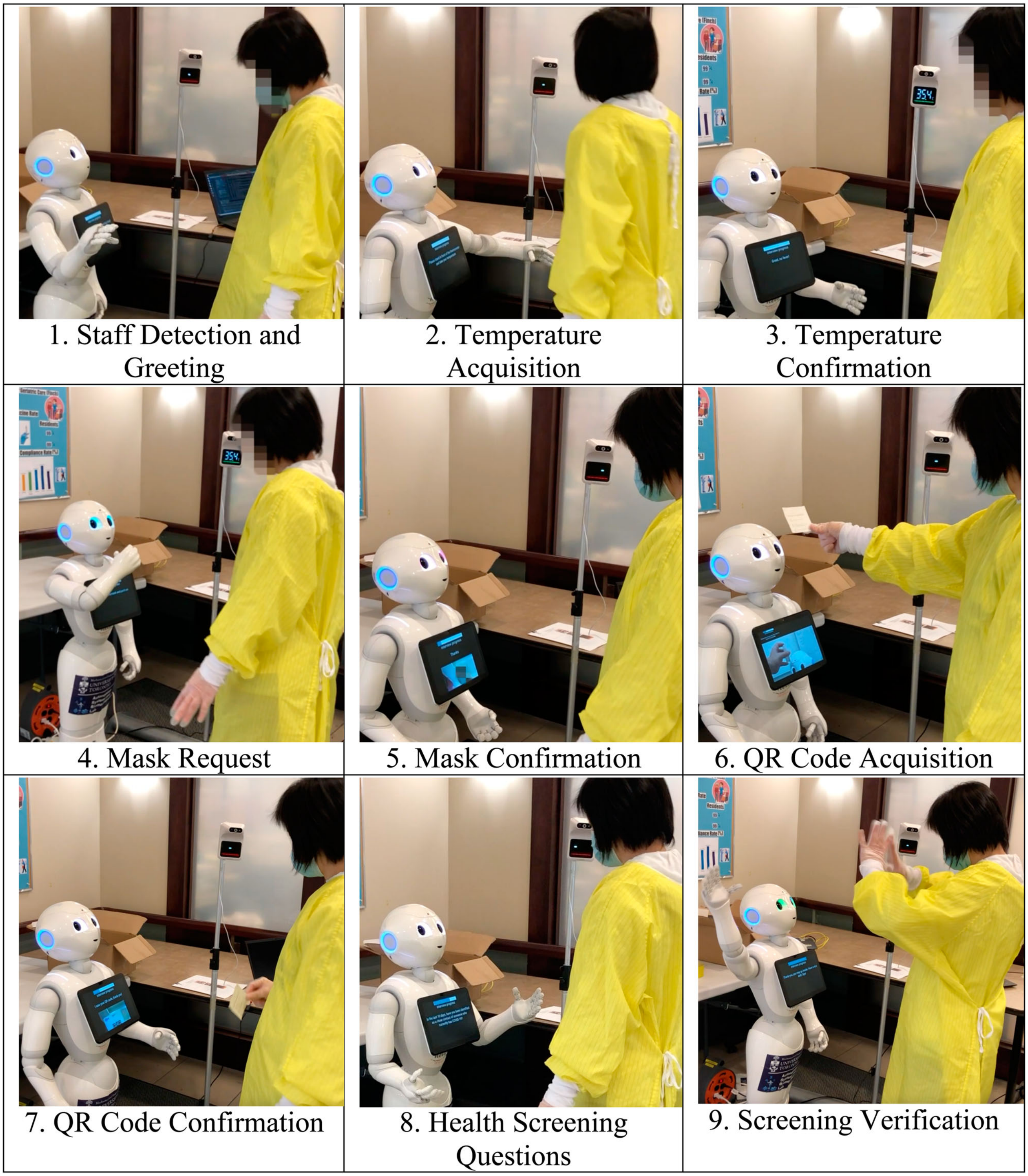


Figure 3. Robot screening images in 9 steps.

questionnaires consisted of 7 main attributes: screening experience, perceived efficiency, cognitive attitude, freeing up staff, perceived safety, affective attitude, and intent to use the robot. Questions for attitude, perceived enjoyment, and intent to use were based on the Almere model [37]. The pre-study questionnaire also contained demographic information (age range, gender, occupation) and previous robot experience (no experience, beginner, intermediate, advanced). In total, we gathered 56 pre-study and 27 post-study questionnaires from the same group of staff participants. The smaller number of post-study questionnaires was due to the prevalence of the Omicron variant of the virus, which placed the LTC home in lockdown and stopped our HRI

Table 1. List of robot behaviors and actions for the 9 screening steps.

Robot Behavior	Robot actions (speech and gestures). Corresponding text for robot's speech is displayed on its tablet.
1. Staff Detection and Greeting	Pepper waves and says, 'Hello! I'm Pepper, the health screening robot.'
2. Temperature Acquisition	Pepper points to the contactless thermometer and says, 'Please take your temperature.'
3. Temperature Confirmation	If no fever is detected ($<$ 37.5°C), Pepper says, 'Great, no fever!' If an elevated temperature is detected ($>$ 37.5°C), Pepper sounds an audible alarm and says, 'Please check with reception.'
4. Mask Request	Pepper gestures with its hand and says, 'Please take a mask and put it on.'
5. Mask Confirmation	Pepper checks to confirm if mask is on correctly (using its forehead RGB camera to take a photo and an adapted version of AIZoo Tech's FaceMaskDetection software [31]) and displays an image of the person on its tablet. If their mask is on properly, Pepper confirms this by saying, 'Thanks!' If their mask is not on properly, Pepper says, 'Is your mask on properly? I can't tell yet.' If there are too many people in the robot's viewfinder, Pepper says, 'One at a time please. Remember to keep social distancing.'
6. QR Code Acquisition	Pepper gestures to its forehead camera and says, 'Please show me your QR code. Hold it in front of my forehead.' QR code scanning is done with the robot's forehead RGB camera and Pepper's barcode reader software.
7. QR Code Confirmation	A confirmation screen displaying an image of the QR code will show on the tablet, and Pepper says, 'Got it, thank you very much.' If no QR code is scanned, Pepper says, 'Please try again.'
8. Health Screening Questions	Pepper asks health screening questions provided by the Ontario Ministry of Health (which are updated regularly), and waits for a Yes/No answer before proceeding to the next question.
9. Screening Verification	If all answers to the screening questions are *No*, Pepper waves and says, 'Thank you, you may go inside. Have a nice shift, bye!' If staff answered *Yes* to one or more questions, Pepper sounds an audible alarm and says, 'Please go see reception.'

Table 2. Descriptive statistics for pre-study, with post-study questionnaire results in parentheses.

	Descriptive Statistics Pre(Post)-Study			
Questions Pre-Study (Post Study)	*Median* ($\tilde{x}$)	*IQR*	*Min*	*Max*
Q1. *Screening experience*				
I have had a good experience with the way the health screening (the robot health screening) is being conducted at Yee Hong	4(4)	2(0)	2(3)	5(5)
Q2. *Efficiency*				
It would be (it is) more efficient if the screening was done/is done automatically/with the robot	4(4)	2(0.75)	1(2)	5(5)
Q3. *Cognitive attitude*				
I think having a robot ask COVID health screening questions would be (is) a good idea	4(5)	1.5(1)	1(1)	5(5)
Q4. *Freeing up staff*				
Using a robot would (did) free up staff that need to do the screening	4(4)	1(1)	1(5)	5(5)
Q5. *Safety*				
I think a robot would make (makes) the health screening process safe	4(5)	1(1)	1(3)	5(5)
Q6. *Affective attitude*				
I think a robot will make (makes) the screening process enjoyable	4(4)	1(0.75)	1(2)	5(5)
Q7. *Intent to use*				
I would (would continue to) use a robot to do the COVID screening at Yee Hong	4(4)	1(0.75)	1(3)	5(5)

study as our researchers were not allowed into the facility at this time.

For administrators and management staff only: An additional, post-study questionnaire for administrators and management staff, who would be in charge of deployment, set-up and adoption of the screening robot within the home, also included the NASA-TLX [18], and the SUS [22,38]. We focused on management and administrative staff within the LTC home, as management plays a key role in implementing robotics, as shown in [39]. The NASA-TLX, listed in Appendix 3, was used to determine how much effort, both mentally and physically, was required to set up and use the robot. The SUS, listed in Appendix 4, was used to measure the usability of the robot for the screening task. Eleven questionnaires were obtained from administrators and managerial staff to directly assess perceived usability and perceived workload.

5. Results

Our HRI study investigated staff members' overall experiences with a social screening robot in a LTC facility and explored how demographic information influenced staff's attitudes and their intent to use a robot [9]. Table 2 presents the pre/post-study questionnaire with descriptive statistics.

We conducted non-parametric tests as we are using Likert scale data, which are ordinal data. Non-parametric tests are recommended to be used when the sampling distribution is non-normally distributed [40]. We conducted a series of Shapiro–Wilk tests of normality, and concluded that our data were non-parametric ($p < 0.05$).

When comparing overall pre- and post-study results using a Wilcoxon Signed Rank (WSR) test, a statistical significance ($Z = 2.060, p = 0.039$) was found for cognitive attitude (Q3) after participants interacted with Pepper ($\tilde{x} = 5$, $IQR = 1$) compared to prior to interaction ($\tilde{x} = 4, IQR = 2$).

Gender: The 56 pre-study questionnaire participants consisted of 31 women, 8 men, and 17 did not specify a gender. The 27 post-study questionnaire participants consisted of 15 women, 11 men, and 1 did not specify a gender. We performed a WSR test to compare within the gender groups before and after the study, and a statistical significance ($Z = 2.000$, $p = 0.046$) was found for cognitive attitude (Q3) for men after interacting with Pepper ($\tilde{x} = 5$, $IQR = 1$) compared to prior to interaction ($\tilde{x} = 4$, $IQR = 0.25$).

Age: In our pre-study questionnaire, the age distribution was 20–60+; the median age group was 40–49. In the post-study, the age distribution was 30–60+; the median age group was 50–59, as there were no participants post-study in the 20–29 age group. No significant differences were found between age groups as determined by Kruskal Wallis (KW) tests ($p > 0.05$), or within age groups when comparing results prior to and after having interacted with the robot, as determined by WSR tests ($p > 0.05$).

Occupation: Staff occupations were grouped into: (1) Administrators (Admin) and Managers (pre-study $n = 11$, post-study $n = 12$), which included those working in human resources, reception, information technology, and management roles; (2) Nurses, including nurse practitioners, registered nurses, and registered practical nurses (pre-study $n = 8$, post-study $n = 4$); (3) Personal Support Workers (PSW) (pre-study $n = 14$, post-study $n = 7$); and (4) Rehabilitation & Social Care (RSC) (pre-study $n = 8$, post-study $n = 4$), including social workers, recreational/activation coordinators, physiotherapists, occupational therapists, and dieticians. No significant differences were found between occupation roles, as confirmed by KW tests. When comparing the same occupation group prior to and after having interacted with the robot, results showed that the RSC group had a statistical significance for Q2 (efficiency), as determined by WSR tests ($Z = 2.000$, $p = 0.046$). The RSC group had a slightly higher median score prior to interacting with the robot ($\tilde{x} = 4$, $IQR = 1$) than after interacting with Pepper ($\tilde{x} = 3.5$, $IQR = 1$). A statistical difference ($Z = 2.000, p = 0.046$) was also found for Q4 (freeing up staff). The RSC group had a slightly higher median score prior to interacting with the robot ($\tilde{x} = 4.5$, $IQR = 1$) than after interacting with Pepper ($\tilde{x} = 4$, $IQR = 0.25$). Results also showed the PSW group had a statistical significance for Q5 (safety), as determined by a WSR test (Z $= 2.070$, $p = 0.038$). This group thought a robot would make the health screening process safer after having interacted with Pepper ($\tilde{x} = 5$, $IQR = 1$) than prior to interaction ($\tilde{x} = 4$, $IQR = 1.75$).

Previous Robot Experience: For Previous Robot Experience, responses included No Experience ($n = 11$); Beginner ($n = 10$), defined as seeing robots on television or at museums; Intermediate ($n = 4$), defined as seeing robots used in the workplace, delivering packages or interacting with residents; and Advanced ($n = 2$), defined as hands-on experience using a robot at work. Differences were found between the groups; namely, results showed a statistically significant difference for Q2 (efficiency), as determined by a KW test ($H(2) = 6.018$, $p = 0.049$). Post-hoc non-parametric MWU tests with Bonferroni correction of $\alpha = 0.016$ ($U = 26.5$, $Z = -2.244$, $p = 0.043$, $r = 0.49$) showed those with no prior robot experience rated robot

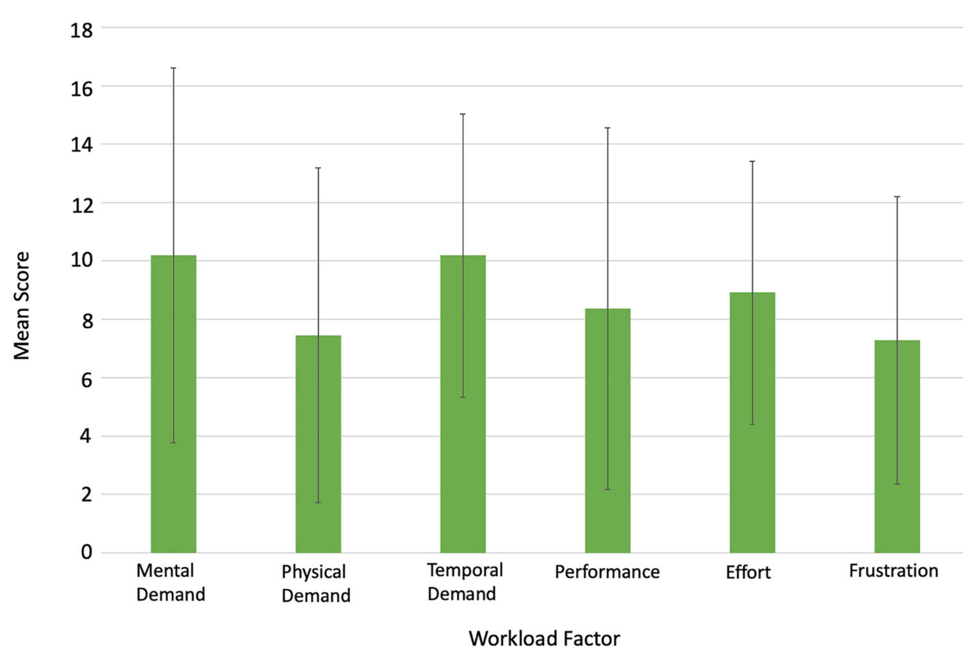


Figure 4. NASA-TLX Mean Scores for Workload Factors, with Standard Deviation Represented by Vertical Lines.

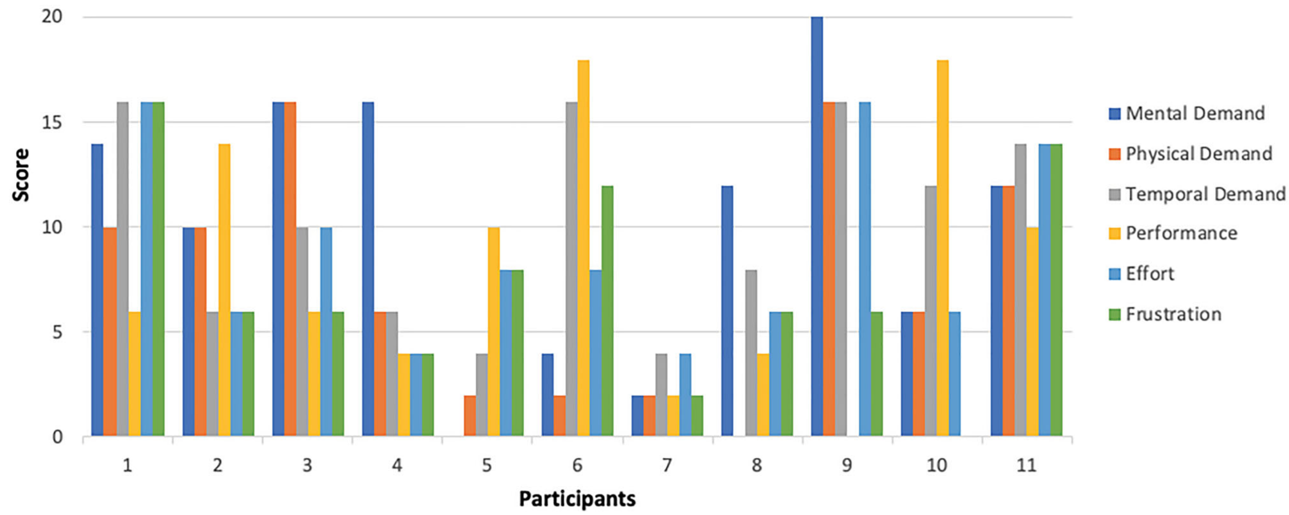


Figure 5. NASA-TLX individual scores for administrator and managerial staff for robot screening task.

efficiency higher ($\tilde{x} = 4$, $IQR = 1$, min $= 4$, max $= 5$) than those with beginner experience ($\tilde{x} = 4$, $IQR = 1$, min $= 2$, max $= 5$). No statistical significance was found between the other experience groups.

To address research questions presented in Section 3, we measured perceived workload using NASA-TLX and perceived usability using SUS.

5.1. Perceived workload

The NASA-TLX mean scores and individual participant scores are presented in Figure 4 and Figure 5. The six workload factors (mental demand, physical demand, temporal demand, performance, effort, frustration) are presented with mean scores (μ) and standard deviations (σ) for each workload factor, along with the overall workload. The individual scores ranged from 10.67 to 70.67. These scores were weighted, by presenting participants with workload factor pairs and asking them to choose the factor that was more important to their experience of workload for the robot screening task. The overall workload was then calculated: $\mu = 44.67$, $\sigma = 19.05$.

5.2. Perceived usability

Each statement of the SUS questionnaire and its corresponding descriptive statistics are presented in Table 3. Individual participant SUS scores are presented in Table 4, and the overall SUS score is presented in Figure 6, with $\mu = 62.5$, $\sigma = 14.5$, $\tilde{x} = 57.5$, min $= 45$, max $= 90$.

Staff rated their willingness to frequently use the robot for screening high (Statement 1, $\tilde{x} = 4$, $IQR = 1$), as well as how quickly they learned to use the robot (Statement 7, $\tilde{x} = 4$, $IQR = 1$). They also felt confident using the robot screener (Statement 9, $\tilde{x} = 4$, $IQR = 1$) and did not find it awkward to use (Statement 8, $\tilde{x} = 2$, $IQR = 2$). Staff rated the remaining questions as neutral. Namely, they were neutral about the robot being too complex for the screening task (Statement 2, $\tilde{x} = 3$, $IQR = 1.5$); that they would need the support of a technical person to be able to use the robot screener (Statement 4, $\tilde{x} = 3$, $IQR = 1$); that there was too much inconsistency with the robot screening system (Statement 6, $\tilde{x} = 3$, $IQR = 1$); and that they needed to learn a lot of things before they could use the robot screener (Statement 10, $\tilde{x} = 3$, $IQR = 3$). They were also neutral in their rating of the robot being easy to use for screening (Statement 3, $\tilde{x} = 3$, $IQR = 1$), and at finding the various functions of the robot screening system well integrated (Statement 5, $\tilde{x} = 3$, $IQR = 1$). The rating for Statement 3 had a 64% frequency for the neutral rating; the remaining 36% rated it positively, which can correspond to the moderately low effort rating on the NASA-TLX.

6. Discussion

Our mean overall workload score falls within the lower quartile of the mean workload scores for robot operation tasks, as reported in [41]. Furthermore, it is also lower than robot teaching tasks performed by care workers via learning from demonstration [15]. This indicates lower workload during autonomous screening on the part of the user. With respect to socially assistive robots, overall workload for care workers was also lower for the screening task compared to the learning from demonstration task.

In general, perceived physical demand was rated low ($\mu = 7.45$, $\sigma = 5.73$). There was little physical demand

Table 3. SUS Questionnaire with median ($\tilde{x}$), and frequency.

Statement	Median ($\tilde{x}$)	*IQR*	Frequency 1	2	3	4	5
1. I think that I would like to use the robot frequently for screening.	4	1	0	1	3	5	2
2*. I found using the robot for screening too complex.	3	1.5	3	2	6	0	0
3. I thought it was easy to use the robot for screening.	3	1	0	0	7	3	1
4*. I think that I would need the support of a technical person who is always nearby to be able to use the robot screener.	3	1	1	2	5	2	1
5. I found the various functions of the robot screening system were well integrated.	3	1	0	0	6	5	0
6*. I thought there was too much inconsistency in the robot screening system.	3	1	1	3	5	2	0
7. I would imagine that most staff would very quickly learn to use the robot screener.	4	1	0	2	2	5	2
8*. I found the robot screener very awkward to use.	2	2	4	3	3	1	0
9. I felt very confident using the robot screener.	4	1	0	1	4	5	1
10*. I needed to learn a lot of things before I could use the robot screener.	3	3	4	0	3	3	1

* Statements are negatively worded.

Table 4. Participant SUS scores.

Participant	1	2	3	4	5	6	7	8	9	10	11
SUS score	57.5	57.5	47.5	52.5	62.5	70.0	80.0	90.0	50.0	75.0	45.0

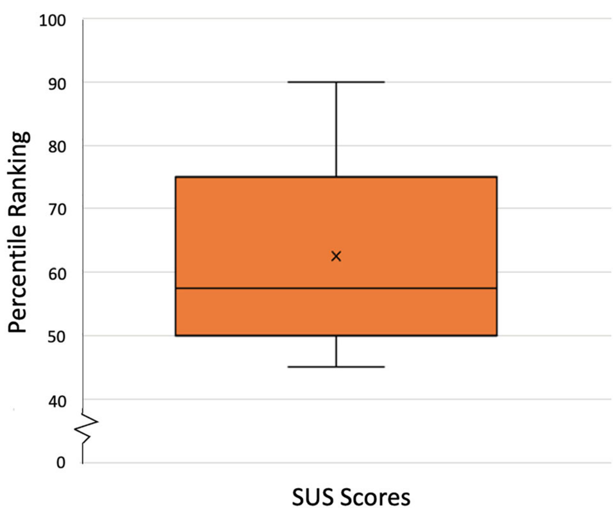


Figure 6. Box and whisker plot of SUS scores with quartiles (box), min-max (whisker), median (line), and mean (*x*).

on the user during the screening process, as the interaction was mostly social, and only required the user to stand in place. The user only needed to turn to face the robot and temperature sensor, hold up their QR code or put on a mask. When interacting with a SAR, physical demand is usually low, since the human is not required to physically manipulate the robot, and interaction is done through verbal and non-verbal communication [15]. The screening robot also interacted in an autonomous multimodal manner. The results also showed both perceived mental ($\mu = 10.18, \sigma = 6.42$) and temporal ($\mu = 10.18$, $\sigma = 4.85$) demand were rated moderate. Mental demand involves mental and perceptual activity, including thinking, deciding, looking, waiting to speak [18]. During the robot screening, the user had to interact with the robot in several ways: through speaking, following the robot's instructions, and waiting for their turn to complete an action. The user needed to follow the robot's instructions in the order they were given; they focused on the robot and paid close attention to its instructions, which is directly linked to mental demand. A similar outcome was found in [15].

Temporal demand is related to the time it takes to complete the screening task and the pace of the screening process [18]. Staff needed to complete the screening process quickly in order to start their shifts on time. Even though the robot screening process itself was short (an average of 80 s), it was observed that many people were in a rush at the beginning of their shifts. They had to be patient, for example, by letting the robot first finish speaking before answering its questions. It is possible that a moderate temporal demand could also be due to unfamiliarity with the robot [15], since a lot of the staff had no direct prior experience with a robot in the workplace.

Overall, staff members rated their Performance as moderately good ($\mu = 8.36$, $\sigma = 6.19$) and Effort as moderately low ($\mu = 8.91$, $\sigma = 4.5$) during screening with Pepper. Moderately good performance is associated with satisfaction and successfully accomplishing the goal of a task [18]. The moderately low effort shows staff were able to follow the robot's screening instructions, answer questions easily, and successfully accomplish the screening task each time. In general, the lower the effort, the higher the usability [42].

It is interesting to note that staff rated Frustration as the lowest workload factor ($\mu = 7.27$, $\sigma = 4.92$). We postulate that since the admin and management staff observed first-hand the benefits of using robots during the pandemic on staff resources, and since frustration can be caused by time delays [43], minor delays due to robot speech recognition, where the robot had to ask a person to repeat what they said due to the noisy environment, were tolerated. Having low frustration is a benefit to conducting an HRI task [43].

The overall mean SUS score ($\mu = 62.5$, $\sigma = 14.5$) falls between the *OK* and *Good* rating, when compared to the adjective rating scale in [28]. Improvements can be made to the robot's speech recognition in noisy and crowded environments. As the majority of staff had no prior experience to beginner experience with robots, this may have impacted the overall SUS score, as there is evidence that has shown SUS is related to users' experience with a technology product [38]. Namely, people with more experience will be more likely to provide more favorable ratings [44].

In general, the robot was found to be easy to use, as staff rated frustration, perceived physical demand and effort low on the NASA-TLX. This addresses RQ1: Interacting with a SAR for screening has minimal impact on staff workload. Additionally, as measured by the frequency of responses on the SUS questionnaire, staff had high willingness to use the robot for screening, and quickly learned to use it. This addresses RQ2: Robot perceived usability impacts the willingness of staff to use the SAR.

A human-like social robot was used for our HRI study to promote natural communication, and eliminate the need to train all the staff when using the robot for the screening task during this high-stress period. Existing systems, such as virtual agents and sensors, could achieve similar screening outcomes. However, socially expressive robots allow for natural face-to-face communication, which provides benefits to user engagement, leading to more successful task completion than when embodied robots are not used [45]. Furthermore, they provide a more engaging and enjoyable interaction as compared to interaction with a non-expressive robots [46]. In a survey of 33 studies comparing human interaction with physical robots and with virtual agents, it was shown that physically present robots are more persuasive and were perceived more positively than virtual agents [47]. For example, when comparing a robot to a tablet delivering healthcare instructions, a robot was rated as more sociable, and users had more positive interactions with the robot, along with increased participation in the suggested activity [48].

Our set-up used an external free-standing non-contact thermometer, as this closely resembles the screening set-up used by the human screener. In addition, we used the commercially available social robot Pepper already in our lab, that did not have a built-in thermometer but that could interact in a multimodal manner (using both speech and text), to expedite our study during the pandemic.

We observed that a human screener could complete the screening task in potentially a shorter time frame than the robot screener, an average of 40–80 s for the human screener compared to an average of 80 s for the robot. Screening time depended on environmental factors, such as noise, and how many people entered at once. As our study was not focused on comparing human and robot screening, which is in line with other research on robot adoption in healthcare settings [15,49,50], we did not quantifiably measure each human screening task to determine descriptive statistics. Based on the data the robot collected, technical failures occurred mainly due to: (1) a noisy environment, which affected speech recognition when answering screening questions, and (2) QR code recognition, as some QR codes were presented at a distance or were only partially shown to the camera, causing the robot to time out. The success rate of speech recognition was 64%; this is similar to other HRI studies conducted in noisy environments [26,51,52]. The success rate for QR code recognition was 84% and for mask detection was 100%. A mask detection failure (either an error in detection or a time out) was never encountered, due to the robot requesting that only one person be in close proximity to the robot.

Staff were already used to the human screening format for a full year prior, whereas with the introduction of a robot screener, it was a new technology they were now interacting with. In our HRI study, we investigated staff perceptions and acceptance of the robot over an extended period of time, and within the context of technology adoption. As people become more familiar with technology, they develop usage skills related to it [53]. Therefore, through prolonged interactions, the robot screening procedure could also decrease the overall screening time. It takes more than two months for this familiarity to happen, which will lead to acceptance beyond the novelty effect of adoption [54].

6.1. Subjective responses

Participants provided additional comments regarding the perceived workload and usability of the screening robot. We performed an inductive thematic analysis approach, where we identified patterns and underlying themes in the comments, which we coded and grouped into the following four main themes: (1) enjoyment; (2) familiarization; (3) technical issues; and (4) personalization. Each theme and its associated comments are presented in Table 5.

Table 5. Identified themes for staff comments.

Themes	Example comments
Enjoyment	'I really enjoyed having Pepper here.' 'Overall the robot was a pleasure to have around.' 'Robot is friendly.'
Familiarization	'It just takes a little bit of time to get used to.' 'Patience is required when using autonomous robots.'
Technical issues	'Noisy scenarios and environment also play a role in the robot's effectiveness.' 'The robot had trouble picking up answers due to background noise.'
Personalization	'Maybe can choose different languages.' 'I like how it follows me with its gaze.'

Participants' comments on enjoyment correspond with previous research on robot use and implementation that has found enjoyment to be a key enabler for use [55]. Comments on familiarization suggest that usability would improve over time; familiarization has been found to positively affect implementation [56], and that over time, the ease of use would improve [55].

Technical problems can be a key barrier to implementation [55], as was relayed in the comments. Loud background noise can negatively affect autonomous speech recognition, as has also been noted in [57], [58]. Additionally, as staff in the LTC home were from Hong Kong, Vietnam or mainland China, and Pepper spoke in English, as requested by the LTC home, there were also comments on personalizing the language of the robot to potentially increase robot usability.

7. Limitations

Our robot screening study was impacted by the additional lockdown of the LTC facility due to the presence of the Omicron variant of the virus, which caused us to stop our study, and limited the number of post-study questionnaires collected. Consequently, we did not have the same number of post-study questionnaires as we did pre-study questionnaires. As the data were gathered anonymously, therefore, it was not possible to identify and match the pre-study questionnaires with those who completed the post-study questionnaires.

In general, the robot was able to accurately detect masks and if people were social distancing, but due to background noise at the front entrance of the facility, it sometimes was not able to detect the Yes or No answers to screening questions, which caused it to repeat itself multiple times (up to 3 times). Even though we provided the staff with onsite training and a manual on using the robot so they would be able to deploy the robot on their own for screening, there was high turnover at the front desk, making it difficult for new staff to set-up the robot to run autonomously if they did not receive the training. In the future, we plan to periodically train front desk staff, especially for when outbreaks occur, so that they do not have to rely on the researchers.

8. Conclusions

Our HRI study with an autonomous screening robot for staff took place in a high-risk environment during the COVID-19 pandemic. We investigated the effects of age, gender, occupation, and previous robot experience, along with staff perceptions before and after interacting with the robot, on 7 attributes: screening experience without and with the robot, perceived efficiency, cognitive attitude, freeing up staff, perceived safety, affective attitude, and intent to use the robot. Overall, staff rated all 7 attributes high. We also measured how the SAR impacts workload and how its usability affects care workers in a LTC facility using the standardized NASA-TLX and SUS. In particular, we investigated management and admin staff perceptions of workload and usability of the SAR, as these staff members are key stakeholders in the introduction and uptake of new technology at care facilities. The NASA-TLX results showed that staff found the robot easy to use; and they rated perceived physical demand and effort low. The SUS results showed high willingness to use the robot in a screening context, and staff's ability to quickly learn to use the robot. Our results motivate the application of socially assistive robots from the staff perspective for repetitive tasks within a long-term care home, an area where technological solutions can help with the increasing work demand on staff. Our study shows that directly interacting with a SAR for a mandatory task during a high-stress time did not notably affect staff workload and that staff remained open to using the technology at the home in the future. This supports the use of SARs to help with the increasing work demand on staff.

Future work will include gathering the perceptions, perceived workload, and perceived usability of the robot from a greater number of staff members, including visitors to the facility, over a longer period of time. This will help familiarize staff with the capabilities of SARs, and help increase the usability of and willingness to use SARs for diverse tasks in LTC facilities.

Acknowledgements

The authors would like to thank Yee Hong Centre for Geriatric Care for their participation in our study. In particular, Matt Snyder, Senior Director, Strategy and Innovation, and Lloyd DelRosario, Executive Director, for their support.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Research supported by Manulife CIFAR Population Health & Well-being Grant Program, AGE-WELL NCE Inc., and the Canada Research Chairs (CRC) Program.

Notes on contributors

Cristina Getson received her B.A.Sc. in Engineering Science (Biomedical major) from the University of Toronto and is currently a Ph.D. Candidate in Human-Robot Interaction in the Autonomous Systems and Biomechatronics Lab, Department of Mechanical and Industrial Engineering at the University of Toronto.

Goldie Nejat is a Full Professor and Associate Chair for Research in the Department of Mechanical and Industrial Engineering at the University of Toronto. She is the Canada Research Chair in Robots for Society and the Founder and Director of the Autonomous Systems and Biomechatronics Laboratory. She is also an Adjunct Scientist at both KITE in the Toronto Rehabilitation Institute at the University Health Network, and the Rotman Research Institute at Baycrest Health Sciences.

References

[1] OECD and World Health Organization. Health at a glance: Asia/Pacific 2020: measuring progress towards universal health coverage. In health at a glance: Asia/Pacific. OECD, 2020. doi:10.1787/26b007cd-en

[2] England K, Azzopardi-Muscat N. Demographic trends and public health in Europe. Eur J Public Health. Oct. 2017;27(suppl_4):9–13. doi:10.1093/eurpub/ckx159

[3] United Nations, Department of Economic and Social Affairs, and Population Division. World population ageing 2020 highlights: living arrangements of older persons. 2020.

[4] Combes SJ-B, Elliott RF, Skåtun D. Hospital staff shortage: the role of the competitiveness of pay of different groups of nursing staff on staff shortage. Appl Econ 2018;50(60):6547–6552. doi:10.1080/00036846.2018.1490000

[5] White EM, Wetle TF, Reddy A, et al. Front-line nursing home staff experiences during the COVID-19 pandemic. J Am Med Dir Assoc Jan. 2021;22(1):199–203. doi:10.1016/j.jamda.2020.11.022

[6] Tavakoli M, Carriere J, Torabi A. Robotics, smart wearable technologies, and autonomous intelligent systems for healthcare during the COVID-19 pandemic: an analysis of the state of the art and future vision. Adv Intell Syst. 2020;2(7):2000071. doi:10.1002/aisy.202000071

[7] Ghafurian M, Ellard C, Dautenhahn K. Social companion robots to reduce isolation: a perception change due to COVID-19. In: Ardito C, Lanzilotti R, Malizia A, Petrie H, Piccinno A, Desolda G, Inkpen K, editors. Human-computer interaction – INTERACT 2021. Lecture Notes in Computer Science, vol. 12933. Cham: Springer International Publishing; 2021. p. 43–63. doi:10.1007/978-3-030-85616-8_4

[8] Aymerich-Franch L, Ferrer I. Liaison, safeguard, and well-being: analyzing the role of social robots during the COVID-19 pandemic. Technol Soc Aug. 2022;70:101993. doi:10.1016/j.techsoc.2022.101993

[9] Getson C, Nejat G. The robot screener will see you now: a socially assistive robot for COVID-19 screening in long-term care homes. 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Aug. 2022. p. 672–677.

[10] Jung M, Lazaro MJS, Yun MH. Evaluation of methodologies and measures on the usability of social robots: a systematic review. Appl Sci. Feb. 2021;11(4):1388. doi:10.3390/app11041388

[11] Parasuraman A. Technology readiness index (Tri): a multiple-item scale to measure readiness to embrace new technologies. J Serv Res. May 2000;2(4):307–320. doi:10.1177/109467050024001

[12] Kim S. Working with robots: human resource development considerations in human–robot interaction. Hum Resour Dev Rev. Mar. 2022;21(1):48–74. doi:10.1177/15344843211068810

[13] Blut M, Wang C. Technology readiness: a meta-analysis of conceptualizations of the construct and its impact on technology usage. J Acad Mark Sci. Jul. 2020;48(4):649–669. doi:10.1007/s11747-019-00680-8

[14] Koh WQ, Felding SA, Budak KB, et al. Barriers and facilitators to the implementation of social robots for older adults and people with dementia: a scoping review. BMC Geriatr 2021;21(1):351. doi:10.1186/s12877-021-02277-9

[15] Louie W-YG, Nejat G. A social robot learning to facilitate an assistive group-based activity from non-expert caregivers. Int J Soc Robot Nov. 2020;12(5):1159–1176. doi:10.1007/s12369-020-00621-4

[16] Rantanen T, Leppälahti T, Porokuokka J, et al. Impacts of a care robotics project on Finnish home care workers' attitudes towards robots. Int J Environ Res Public Health. Sep. 2020;17(19):7176. doi:10.3390/ijerph17197176

[17] Broadbent E, Kerse N, Peri K, et al. Benefits and problems of health-care robots in aged care settings: a comparison trial: health-care robots in retirement village. Australas J Ageing. Mar. 2016;35(1):23–29. doi:10.1111/ajag.12190

[18] Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Advances in psychology. Los Angeles (CA): Elsevier; 1988. p. 139–183. doi:10.1016/S0166-4115(08)62386-9

[19] Carros F, Schwaninger I, Preussner A, et al. Care workers making use of robots: results of a three-month study on human-robot interaction within a care home. In: Barbosa S, Lampe C, editors. CHI Conference on Human Factors in Computing Systems. New Orleans (LA): ACM; 2022. p. 1–15. doi:10.1145/3491102.3517435

[20] Papadopoulos I, Ali S, Papadopoulos C, et al. A qualitative exploration of care homes workers' views and training needs in relation to the use of socially assistive humanoid robots in their workplace. Int J Older People Nurs. May 2022;17(3):e12432. doi:10.1111/opn.12432

[21] Papadopoulos C, Castro N, Nigath A, et al. The CARESSES randomised controlled trial: exploring the health-related impact of culturally competent artificial intelligence embedded into socially assistive robots and tested in older adult care homes. Int J Soc Robot. 2022;14:245–256. doi:10.1007/s12369-021-00781-x

[22] Brooke J. SUS – A quick and dirty usability scale, p. 7.
[23] ISO 9241-11:2018(en), ergonomics of human-system interaction — Part 11: usability: definitions and concepts [cited 2022 Dec 18]. Available from: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.
[24] van Greunen D. User experience for social human-robot interactions. In: 2019 Amity International Conference on Artificial Intelligence (AICAI). Dubai: IEEE; 2019. p. 32–36. doi:10.1109/AICAI.2019.8701332
[25] Hornbæk K. Current practice in measuring usability: challenges to usability studies and research. Int J Hum-Comput Stud. Feb. 2006;64(2):79–102. doi:10.1016/j.ijhcs.2005.06.002
[26] Olde Keizer RACM, van Velsen L, Moncharmont M, et al. Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges. Health Technol. Aug. 2019;9(4):595–605. doi:10.1007/s12553-019-00320-9
[27] Huisman C, Kort H. Two-year use of care robot zora in Dutch nursing homes: an evaluation study. Healthcare. 2019;7(1):31. doi:10.3390/healthcare7010031
[28] Bangor A. Determining what individual SUS scores mean: adding an adjective rating scale. 2009;4(3):10.
[29] Kiselev A, Loutfi A. Using a mental workload index as a measure of usability of a user interface for social robotic telepresence. IEEE Int Symp Robot Hum Interact Commun. 2012:4:7A4–7A4.
[30] Iizuka T, Lee YS, Eggleston K. Robots and labour in the service sector. CEPR [cited 2022 Dec 13]. Available from: https://cepr.org/voxeu/columns/robots-and-labour-service-sector.
[31] AIZOO. FaceMaskDetection. Feb. 15, 2022 [cited 2022 Feb 16]. Available from: https://github.com/AIZOOTech/FaceMaskDetection.
[32] WIDER FACE: a face detection benchmark [cited 2022 Mar 9]. Available from: http://shuoyang1213.me/WIDERFACE/.
[33] Ge S, Li J, Ye Q, et al. Detecting masked faces in the wild with LLE-CNNs, p. 9.
[34] Pepper Face Mask detection. SoftBank Robotics Labs, August 7, 2023 [cited 2023 Aug 23]. Available from: https://github.com/softbankrobotics-labs/pepper-mask-detection
[35] Sommana B, Watchareeruetai U, Ganguly A, et al. Development of a face mask detection pipeline for mask-wearing monitoring in the era of the COVID-19 pandemic: a modular approach. In: 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE). Bangkok: IEEE; 2022. p. 1–6. doi:10.1109/JCSSE54890.2022.9836283
[36] Fan X, Jiang M, Yan H. A deep learning based light-weight face mask detector with residual context attention and Gaussian heatmap to fight against COVID-19. IEEE Access. 2021;9:96964–96974. doi:10.1109/ACCESS.2021.3095191
[37] Heerink M, Kröse B, Evers V, et al. Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robot. 2010;2(4):361–375. doi:10.1007/s12369-010-0068-5
[38] Sauro J, Lewis J. Quantifying the user experience – practical statistics for user research. 2nd ed. Cambridge (MA): Morgan Kaufmann; 2016.
[39] Coco K, Kangasniemi M, Rantanen T. Care personnel's attitudes and fears toward care robots in elderly care: a comparison of data from the care personnel in Finland and Japan. J Nurs Scholarsh. 2018;50(6):634–644. doi:10.1111/jnu.12435
[40] Field A, Miles J, Field Z. Discovering statistics using R. London: Sage; 2012.
[41] Grier RA. How high is high? A meta-analysis of NASA-TLX global workload scores. Proc Hum Factors Ergon Soc Annu Meet. 2015;59(1):1727–1731. doi:10.1177/1541931215591373
[42] Komischke T. 7 usability metrics to assess ease of use. CMSWire.com [cited 2022 Dec 20]. Available from: https://www.cmswire.com/digital-experience/usability-testing-7-metrics-to-assess-ease-of-use/.
[43] Weidemann A, Rußwinkel N. The role of frustration in human–robot interaction – what is needed for a successful collaboration? Front Psychol. 2021;12:640186. doi:10.3389/fpsyg.2021.640186
[44] Borsci S, Federici S, Bacci S, et al. Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. Int J Hum-Comput Interact. 2015;31(8):484–495. doi:10.1080/10447318.2015.1064648
[45] Hossen Mamode HZ, Bremner P, Pipe AG, et al. Cooperative tabletop working for humans and humanoid robots: group interaction with an avatar. In: 2013 IEEE International Conference on Robotics and Automation. Karlsruhe: IEEE; 2013. p. 184–190. doi:10.1109/ICRA.2013.6630574
[46] Adalgeirsson SO, Breazeal C. MeBot: a robotic platform for socially embodied telepresence. In: Hinds P, Ishiguro H, editors. 2010 5th ACM/IEEE International Conference on Human-robot Interaction (HRI). Osaka: IEEE; 2010. p. 15–22. doi:10.1109/HRI.2010.5453272
[47] Li J. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum-Comput Stud. 2015;77:23–37. doi:10.1016/j.ijhcs.2015.01.001
[48] Mann JA, MacDonald BA, Kuo I-H, et al. People respond better to robots than computer tablets delivering healthcare instructions. Comput Hum Behav. 2015;43:112–117. doi:10.1016/j.chb.2014.10.029
[49] Ahn HS, Lee MH, MacDonald BA. Healthcare robot systems for a hospital environment: CareBot and ReceptionBot. In: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). Kobe: IEEE; 2015. p. 571–576. doi:10.1109/ROMAN.2015.7333621
[50] Mišeikis J, Caroni P, Duchamp P, et al. Lio-A personal robot assistant for human-robot interaction and care applications. IEEE Robot Autom Lett. Oct. 2020;5(4):5339–5346. doi:10.1109/LRA.2020.3007462
[51] de Jong M, Zhang K, Roth AM, et al. Towards a robust interactive and learning social robot. Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems. AAMAS 2018; 2018.
[52] Kennedy J, Lemaignan S, Montassier C, et al. Child speech recognition in human-robot interaction: evaluations and recommendations. In: Mutlu B, Tscheligi M, Weiss A, et al., editors. Proceedings of the 2017 ACM/IEEE International Conference on Human-robot Interaction.

Vienna: ACM; 2017. p. 82–90. doi:10.1145/2909824.3020229

[53] Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci. Feb. 2000;46(2):186–204. doi:10.1287/mnsc.46.2.186.11926

[54] Sung J, Christensen HI, Grinter RE. Robots in the wild: understanding long-term use. In: Scheutz M, Michaud F, Hinds P, et al., editors. Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction. La Jolla (CA): ACM; 2009. p. 45–52. doi:10.1145/1514095.1514106

[55] Papadopoulos I, Koulouglioti C, Lazzarino R, et al. Enablers and barriers to the implementation of socially assistive humanoid robots in health and social care: a systematic review. BMJ Open. Jan. 2020;10(1):e033096. doi:10.1136/bmjopen-2019-033096

[56] Wu Y-H, Wrobel J, Cornuet M, et al. Acceptance of an assistive robot in older adults: a mixed-method study of human–robot interaction over a 1-month period in the Living Lab setting. Clin Interv Aging. May 2014;9:801–811. doi:10.2147/CIA.S56435.

[57] Mubin O, Henderson J, Bartneck C. You just do not understand me! Speech recognition in human robot interaction. In: The 23rd IEEE International Symposium on Robot and Human Interactive Communication. Edinburgh: IEEE; 2014. p. 637–642. doi:10.1109/ROMAN.2014.6926324

[58] Casey D, Barrett E, Kovacic T, et al. The perceptions of people with dementia and key stakeholders regarding the use and impact of the social robot MARIO. Int J Environ Res Public Health. Nov. 2020;17(22):8621. doi:10.3390/ijerph17228621

Appendices

Appendix 1. Health screening questions

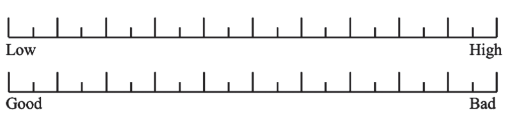
Q1	Do you have a fever, chills, cough, shortness of breath, decreased or loss of taste or smell, or unexplained fatigue?
Q2	In the last 14 days, have you or someone you live with traveled outside of Canada?
Q3	In the last 10 days, have you been identified as a close contact of someone who currently has COVID-19?
Q4	Is anyone you live with experiencing COVID-19 symptoms or waiting for test results after experiencing symptoms?

Appendix 2. Pre/Post study questionnaire with the 7 attributes

Strongly disagree 1	Somewhat disagree 2	Neutral 3	Somewhat agree 4	Strongly agree 5
Q1 (screening experience)		I have had a good experience with the way the health screening (the robot health screening) is being conducted at Yee Hong		
Q2 (efficiency)		It would be (it is) more efficient if the screening was done (is done) automatically with the robot		
Q3 (cognitive attitude)		I think having a robot ask COVID-19 health screening questions would be (is) a good idea		
Q4 (freeing up staff)		Using a robot would (did) free up staff that need to do the screening		
Q5 (safety)		I think a robot would make (makes) the health screening process safe		
Q6 (affective attitude)		I think a robot will make (makes) the screening process enjoyable		
Q7 (intent to use)		I would (would continue to) use a robot to do the COVID-19 screening at Yee Hong		

Appendix 3. NASA-TLX task load index

Please place an 'X' along each scale at the point that best indicates your experience during the robot teaching session, ranging from low to high for statements 1–5 and good to bad for statement 6.



1. *Mental Demand*: How much mental and perceptual activity was required during the robot screening (such as thinking, deciding, remembering, looking, waiting to speak)? For example, was the screening task easy or demanding, simple or complex, forgiving or exacting?
2. *Physical Demand*: How much physical activity was required during the robot screening (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the screening task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
3. *Temporal Demand*: How much time pressure did you feel based on the rate or pace of the robot screening task? Was the pace slow and leisurely or rapid and frantic?
4. Effort: How hard did you have to work (mentally or physically) to accomplish the robot screening task?
5. *Frustration*: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during the robot screening task?
6. *Performance*: How successful do you think you were in accomplishing the robot screening task? How satisfied were you with your performance in accomplishing this task with the robot?

Appendix 4. System usability scale

Please indicate your level of agreement (from Strongly Disagree to Strongly Agree) to the following 10 statements.

1. I think that I would like to use the robot frequently for screening.
2. I found using the robot for screening too complex.
3. I thought it was easy to use the robot for screening.
4. I think that I would need the support of a technical person who is always nearby to be able to use the robot screener.
5. I found the various functions of the robot screening system were well integrated.
6. I thought there was too much inconsistency in the robot screening system.
7. I would imagine that most staff would very quickly learn to use the robot screener.
8. I found the robot screener very awkward to use.
9. I felt very confident using the robot screener.
10. I needed to learn a lot of things before I could use the robot screener.