



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *9th International Conference on Social Robotics (ICSR 2017), Tsukuba, Japan, November 22-24, 2017*.

Citation for the original published paper:

Akalin, N., Kiselev, A., Kristoffersson, A., Loutfi, A. (2017)

An Evaluation Tool of the Effect of Robots in Eldercare on the Sense of Safety and Security

In: Kheddar, A.; Yoshida, E.; Ge, S.S.; Suzuki, K.; Cabibihan, J-J.; Eyssel, F.; He, H. (ed.), *Social Robotics: 9th International Conference, ICSR 2017, Tsukuba, Japan, November 22-24, 2017, Proceedings* (pp. 628-637). Springer International Publishing Lecture Notes in Computer Science

https://doi.org/10.1007/978-3-319-70022-9_62

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-62768>

An Evaluation Tool of the Effect of Robots in Eldercare on the Sense of Safety and Security

Neziha Akalin, Andrey Kiselev, Annica Kristoffersson, and Amy Loutfi

Örebro University, SE-701 82 Örebro, Sweden,
neziha.akalin@oru.se,
<http://mpi.aass.oru.se>

Abstract The aim of the study presented in this paper is to develop a quantitative evaluation tool of the sense of safety and security for robots in eldercare. By investigating the literature on measurement of safety and security in human-robot interaction, we propose new evaluation tools. These tools are semantic differential scale questionnaires. In experimental validation, we used the Pepper robot, programmed in the way to exhibit social behaviors, and constructed four experimental conditions varying the degree of the robot’s non-verbal behaviors from no gestures at all to full head and hand movements. The experimental results suggest that both questionnaires (for the sense of safety and the sense of security) have good internal consistency.

Keywords: sense of safety, sense of security, eldercare, video-based evaluation, quantitative evaluation tool

1 Introduction

To facilitate successful interaction between humans and robots, interaction needs to be natural and share similarities with human-human interaction. Elderly people’s sense of safety and security can increase their quality of life and well-being [9]. Safety, including security, is stated as one of the fundamental needs of human beings [16].

There is no general consensus for the concepts of “safety” and “security” in the literature. One of the generally accepted approaches to distinguish the terms is the intentionality of being caused the harm. The term “safety” refers to unintentional harm (e.g., accidents, careless behavior such as car crashes, occupational injuries, and food poisoning) and “security” refers to intentional harm (e.g., military occupation and computer viruses) [4]. Another approach to differentiate safety and security is the distinction between a system and its environment. Safety refers to hazards the system may cause and harm its environment, security refers to the threats from the system’s environment to negatively affect the system [13,18].

In this paper, we propose a video-based methodology to evaluate the sense of safety and security of a robot intended for scenarios in eldercare. In video-based human-robot interaction (VHRI) methodology, interactive robot scenes

are recorded as movies and shown to the users. VHRI has several advantages like allowing experiments to reach a large number of participants, replacing videos easily according to the participants' ideas and comments, and using the same methodologies in each video including the same robot behaviors, and exact trial instructions [7].

The factors that can influence sense of safety and security can be divided into three different areas:

1. environmental related factors including task complexity and properties of the physical environment,
2. human related factors including personality, emotions, health and mobility, and
3. robot related factors including its behaviors/gestures, performance and autonomy but also its size and shape.

The aim of the study presented in this paper was to develop a quantitative evaluation tool of the sense of the safety and security for robots in eldercare. By investigating the human-robot interaction (HRI) literature and sense of safety and security for elderly people [2,8,11], we propose using semantic differential scale questionnaires.

In the remainder of this paper, an overview of related studies is given in Section 2, the method, experimental design and procedure are described in Section 3. Experiment results are presented in Section 4, and the paper is concluded in Section 5.

2 Related Work

There is a considerable number of contributions dealing with physical safety of robots in the literature [3,10,22]. Nevertheless, there is only a small number of research in HRI that considers sense of safety. The terms that convey a similar meaning as sense of safety are perceived safety [2], psychological safety [14] and mental safety [17].

In [14], Lasota et al. introduced an approach for providing comfortable and stress-free HRI by using robot joint angles and human localization to adjust the speed of the robot. In the paper [14], they defined psychological safety as: *Ensuring that human-robot interaction does not cause excessive stress and discomfort for extended periods of time*. In [12], Kamide et al. also used the term psychological safety and present a new scale to measure safety quantitatively by studying important factors for determining the psychological safety of humanoids. They showed movies of 11 humanoids and asked to the participants open-ended questions about safety. Kamide et al. [12] categorized and analyzed the results and came up with six factors for measuring psychological safety of humanoids. The factors were as follows: performance, acceptance, harmlessness, toughness, humanness and agency, in which the former four factors are more important than the latter two.

Mental safety is defined as not feeling fear, surprise, unpleasant and disgusted while interacting with a robot [17]. The experiments were conducted using virtual robots of varying shape, size and motions. While the emotions fear and surprise are related to sense of security; disgust and unpleasantness are related to comfort according to [17]. Through questionnaires evaluating surprise, fear, disgust and unpleasantness, they observed that robots’ human-like behaviors made the humans more comfortable.

In 2009, Bartneck et al. proposed a series of questionnaires to measure the key concepts in HRI, including the perceived safety [2]. They defined the perceived safety as follows: *Perceived safety describes the user’s perception of the level of danger when interacting with a robot, and the user’s level of comfort during the interaction.* Their Godspeed V: Perceived safety questionnaire is commonly used in different HRI scenarios. For example, [15] used the Godspeed V questionnaire to compare two different navigation algorithms to investigate the effect of legibility on the perceived safety in a path crossing scenario. Weiss and Bartneck presented a meta-analysis of the usage of the Godspeed Questionnaire Series in HRI studies in 2015 [21]. They found that the perceived safety questionnaire had been used in 37 different studies.

In [20], they presented a survey on security robots by examining 60 projects. They investigated crime-fighting robots, their development and applications into four classes which were teleoperated, distributed, surveillance, and law-enforcement robot architectures. On the one hand, the robot security studies elaborated using the robots for crime-fighting, but on the other hand, the conceptualization of the sense of security is not yet fully handled. In [17], mental safety and physical safety was defined despite using sense of security in the publication’s title.

3 Methods

To create more accurate human-robot real world scenarios, nonverbal gestures like head nodding, gaze, iconic and metaphoric gestures were included in the scenarios. To gain a better understanding of how these gestures affect the sense of safety and security, we conducted between-subjects video-based experiments using the Pepper humanoid robot. The current study considered four different scenarios comprising daily life activities in which elderly people may be engaged. The participants’ sense of safety and security was measured through different variations of nonverbal gestures including only arms gestures (configuration 1), only head gestures (configuration 2), head and arms gestures (configuration 3) and no gestures (configuration 4). The scenarios used in this study were recorded as short videos which were between 20 seconds and 30 seconds long. We have selected four scenarios from [6] which presented a video based evaluation to compare elderly people’s perceptions of socially assistive domestic robots in two different cultural backgrounds: Italian and Swedish user groups. Two of the selected scenarios are proactive (i.e. the robot is the initiator) and two of them

are on-demand scenarios (i.e. the user is the initiator). Detailed explanations about the scenarios are given in Section 3.1.

The robot used in our experiments was Pepper. The Pepper is a humanoid robot with 20 degrees of freedom, has a height of 1.2 m and a weight of 29 kg. It has 3 multi-directional wheels and is equipped with two cameras and one 3D camera, 4 directional microphones, 5 touch sensors and a loudspeaker[1].

3.1 Videos

Scenario number 1 and 2 represent proactive situations; scenario number 3 and 4 are on-demand interactions. The selected scenarios, which are taken from [6], are shortly described below:

Scenario 1: Environmental safety. *The actor is sitting on the sofa, watching TV. Meanwhile, in the kitchen, the sauce on the stove is overcooking. The robot moves towards the actor and says: “The pot is burning. You should turn it off.” The actor immediately goes to the kitchen and turns the stove off.* (1a)

Scenario 2: Reminding analysis. *The actor is in the kitchen. He is about to have breakfast. When he puts the pot on the stove, the robot says: “You cannot have breakfast now. You have an appointment for a medical analysis.” The actor answers: “You’re right. I have forgotten all about it!”* (2a)

Scenario 3: Finding objects. *The actor is sitting on the sofa, and takes a magazine to read. Suddenly, he realizes that the glasses are not on the table in front of him. The actor calls the robot and asks: “Where are my glasses?” The robot answers “Just a minute, I am checking” and then the robot answers: “The glasses are in the kitchen.” The actor goes to the kitchen and takes the glasses.* (1b)

Scenario 4: Reminding medication. *The actor is sleeping on the sofa, and suddenly wakes up. He does not realize what time it is, and thus he asks the robot. The robot answers: “It is four o’clock.” The actor does not remember whether or not he took his medicine after lunch, and asks the robot. The robot answers: “Yes, you took it.”* (2b)

Figure 1 and 2 shows screenshots from each scenario.

3.2 Participants

100 participants (47 males and 53 females) in ages ranging from 14 to 62 years ($M=35.48$, $SD=10.58$) took part in the study. The advertisement of the on-line survey was done in social media and through mailing lists. For each of the four conditions, 25 volunteer participants filled out the survey. The participants’ familiarity with robots is as follows: 27 of them had previously seen a real robot but were not familiar, 20 of them had already interacted with robots, 22 of them worked/work with robots and 31 of them had previously seen a robot on TV/Internet.

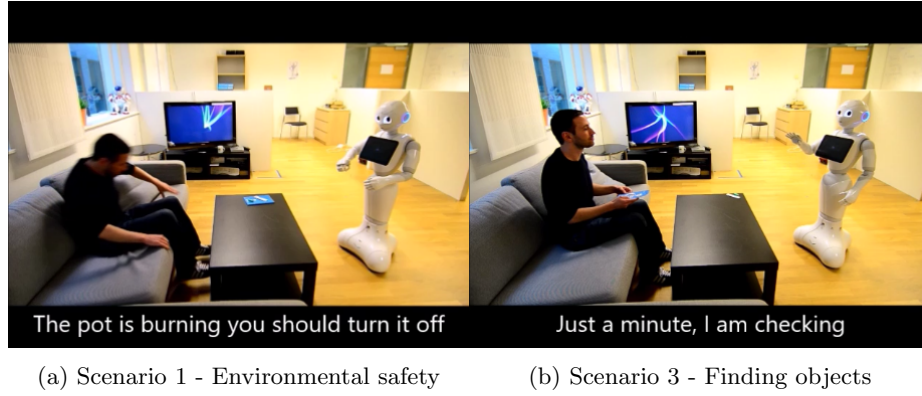


Figure 1: Screenshots from the scenarios (a) Scenario 1 - Environmental safety, the robot has both arm and head gestures (configuration 3) , (b) Scenario 3 - Finding objects, the robot has only arm gestures (configuration 1)

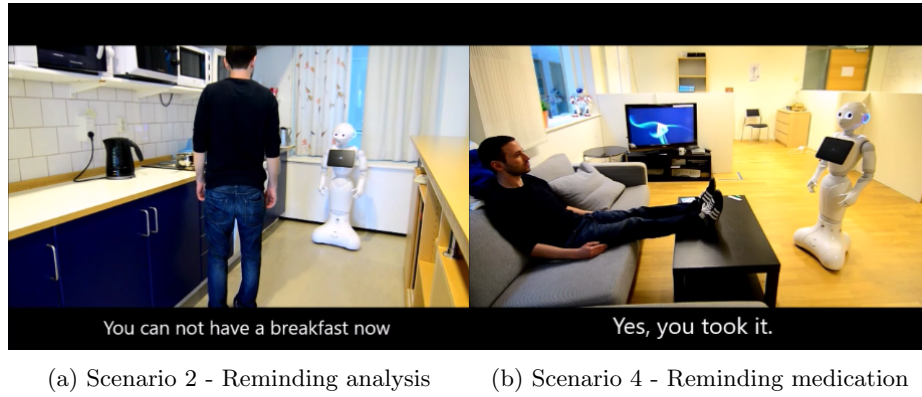


Figure 2: Screenshots from the scenarios (a) Scenario 2 - Reminding analysis, the robot has only head gestures (configuration 2) , (b) Scenario 4 - Reminding medication, the robot has no gestures (configuration 4)

3.3 Questionnaires

The questionnaires were prepared based on a literature study given in Section 2. The terms relating to sense of safety and sense of security were collected and presented as differential scale questionnaires.

The questionnaires used for data collection consisted of one part for socio-demographics plus four separate sections. In the first part, the participants were asked about gender, age, educational level, country of residence and their familiarity with robots. The brief explanation about the rest of the questionnaires is given below:

Section 1. Six questions were designed in a five point semantic differential scale to assess sense of safety with regards to the four videos that were presented. Three of them were taken from Godspeed V [2]. The questionnaire includes the perceived safety questionnaire from the Godspeed questionnaire series, adding the following items: threatening – safe, uncomfortable – comfortable and predictable – unpredictable (all of them on a 5-point semantic differential scale).

Section 2. Six semantic differential scale questions to assess sense of security in response to the four videos that were presented. We designed a questionnaire for the sense of security including the items: insecure – secure, unfamiliar – familiar, fear – ease, unreliable – reliable, unnatural – natural, lacking control – in control.

Section 3. Seven questions rated on a five-point likert scale, ranging from “Strongly disagree” to “Strongly agree” to assess the acceptability of the robot. Four of them were asked after each video and three of them were asked at the end of the survey. These questions were taken from [8] and [11].

Section 4. Self-Assessment Manikin (SAM) [5] was used to evaluate the participants’ emotions. Manikin is a scale for assessing emotions in a valence and an arousal scale which ranges from unpleasant to pleasant on the valence scale and excited to calm on the arousal scale. At the end of the survey, an additional text field was provided allowing for free comments. The questions are given in Table 1.

Table 1: The questions asked throughout the survey.

Measure	Questionnaire items	Scale
Sense of safety	Threatening - safe, anxious - relaxed, agitated - calm, quiescent - surprised, uncomfortable - comfortable, unpredictable - predictable	1 to 5
Sense of security	Insecure - secure, unfamiliar - familiar, fear - ease, unreliable - reliable, unnatural - natural, lacking control - in control	1 to 5
Acceptance	I think the robot would be helpful in my home I would trust the robot if it gave me advice I would follow the advice the robot gives me	1 = Strongly disagree 5 = Strongly agree
Scenarios	The scenario seems realistic, one can encounter this scenario in the daily life. I liked the scenario I can use the robot for a longer period in my home I can imagine having a robot taking care of me	1 = Strongly disagree 5 = Strongly agree
Emotions	Unpleasant - pleasant, excited - calm	1 to 9

3.4 Procedure

The data was collected using an online survey including videos of the robot and elderly interactions. It is in English and can be found at the link <https://goo.gl/forms/94db0D96yR5nYPhl2>. The participants were provided with

a short introduction about the survey, and filled out the socio-demographics part. Then, the video sessions started. The participants were asked to watch the videos one by one. At the end of each video, a short questionnaire was used to assess the subjects' views on the scenario and the robot. After the videos, the participants completed four questionnaires to assess their sense of safety, sense of security, emotions and acceptance. No explanation about the real aim of the experiments was given in order to avoid influencing and biasing the participants.

4 Experimental Results

To compare the effects of nonverbal gestures on the sense of safety, sense of security and acceptance one-way Analysis of Variance (ANOVA) was applied. Results seem to indicate that the participants did not notice the gestures of the robot. Throughout the survey, we did not mention anything about the gestures. We expected that the gestures would affect the participants, but the results show that there is no statistically significance between the four configurations for all the measures. The reason for this could be that the participants concentrated on the scenarios and the idea of using robots at homes. Only a few participants noticed the gestures and made comments about them. The participants also rated their emotions using SAM that provided us to place emotional states onto Russell's valence-arousal scale [19]. Positive emotions like happy, pleased, delightful etc. are placed in the first quadrant (high valence, high arousal). For all configurations, the average of participants rated emotional state were in the first quadrant.

4.1 Sense of Safety

We conducted a reliability analysis to check the internal consistencies within the items in the questionnaires. The calculated Cronbachs alphas for all participants and configurations exceed 0.7. This indicates that the questionnaire items had a good internal consistency. The Cronbach's alpha for the sense of safety questionnaire including all items is 0.787. Most of the participants were confused about the quiescent – surprised scale so we excluded this item and recalculated the Cronbach's alpha. The final Cronbach's alpha value was 0.868. The internal reliabilities for each configuration were also calculated and given in Table 2 where New Cronbach's alpha shows the values after the quiescent – surprised item is excluded. The likert scale descriptive statistics (mean, median and standard deviation) for each configuration can be seen in Table 3. Principal Component Analysis (PCA) and Factor Analysis (FA) was also applied on the questionnaire items. PCA showed that the most important items in the sense of safety questionnaire were threatening – safe, anxious – relaxed and agitated – calm. FA showed that there were two factors and all the items are related to the same factor, except for quiescent – surprised.

Table 2: Sense of safety Cronbach's alpha values for each configuration.

Configuration	Cronbach's α	New Cronbach's α
c1	0.784	0.874
c2	0.880	0.889
c3	0.863	0.934
c4	0.810	0.856

Table 3: Sense of safety likert scale descriptive statistics for each configuration.

Configuration	Mean	Median	Standard deviation
c1	3.83	4	0.95
c2	3.67	4	0.96
c3	3.84	4	1.22
c4	3.70	4	1.02

4.2 Sense of Security

The Cronbach's alpha value for all participants and all configurations was 0.870 which shows the questionnaire has good internal consistency. The internal reliabilities for each configuration are given in Table 4. The likert scale descriptive statistics of sense of security (mean, median and standard deviation) for each configuration can be seen in Table 5. PCA analysis showed that the most important items in the sense of security questionnaire are insecure – secure, unfamiliar – familiar, fear – ease, unreliable – reliable. FA showed that there are two factors in the questionnaire.

Table 4: Sense of security Cronbach's alpha values for each configuration.

Configuration	Cronbach's α
c1	0.854
c2	0.884
c3	0.929
c4	0.826

Table 5: Sense of security likert scale descriptive statistics for each configuration.

Configuration	Mean	Median	Standard deviation
c1	3.62	4	0.97
c2	3.72	4	1.00
c3	3.72	4	1.24
c4	3.49	4	1.10

5 Conclusions and Future Work

In this paper, we proposed a questionnaire-based framework for evaluating sense of safety and sense of security in eldercare scenarios. Four experimental conditions with varying the degrees of the robot’s non-verbal behaviors ranging from no gestures at all to full head and hand movements were used to test the questionnaire. We prepared videos showing proactive and on-demand situations, and used our evaluation framework to investigate whether such small differences in the robots behaviors (varying degree of non-verbal behavior) would make a statistically significant difference on how the interaction scenarios are perceived in terms of sense of safety and sense of security. In total, 100 subjects participated in the experiment. The experimental results suggested that both questionnaires (for the sense of safety and the sense of security) have good internal consistency. Particularly, we obtained Cronbach’s alpha of 0.868 and 0.870 for the sense of safety and sense of security respectively. Additionally, factor analysis showed that all the items correlated with each other for both questionnaires.

Interestingly, evaluating the specific effects of non-verbal behavior on the perceived sense of safety and sense of security, the experiment did not reveal any statistical difference between the conditions. Understanding the reasoning behind this would clearly require a much larger study in which other robots and/or variations in functionality could be evaluated and compared. Thus, the primary direction of our future work is to implement an experimental design that will reveal what characteristics of robot systems affect the sense of safety and security. For now, the questionnaire and video-based evaluation presented in this work, are the first steps towards this larger goal.

References

1. Pepper. <https://www.ald.softbankrobotics.com/en/cool-robots/pepper>. Accessed: 2017-06-30.
2. Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.
3. Antonio Bicchi, Michael A Peshkin, and J Edward Colgate. Safety for physical human–robot interaction. In *Springer handbook of robotics*, pages 1335–1348. Springer, 2008.
4. Max Boholm, Niklas Möller, and Sven Ove Hansson. The concepts of risk, safety, and security: applications in everyday language. *Risk analysis*, 2015.
5. Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.
6. Gabriella Cortellessa, Massimiliano Scopelliti, Lorenza Tiberio, Gion Koch Svedberg, Amy Loutfi, and Federico Pecora. A cross-cultural evaluation of domestic assistive robots. In *AAAI Fall Symposium: AI in Eldercare: New Solutions to Old Problems*, pages 24–31, 2008.

7. Walters et al. Evaluating the robot personality and verbal behavior of domestic robots using video-based studies. *Advanced Robotics*, 25(18):2233–2254, 2011.
8. David Fischinger, Peter Einramhof, Konstantinos Papoutsakis, Walter Wohlking, Peter Mayer, Paul Panek, Stefan Hofmann, Tobias Koertner, Astrid Weiss, Antonis Argyros, et al. Hobbit, a care robot supporting independent living at home: First prototype and lessons learned. *Robotics and Autonomous Systems*, 75:60–78, 2016.
9. Edit Fonad, Tarja-Brita Robins Wahlin, Kristina Heikkilä, and Azita Emami. Moving to and living in a retirement home: Focusing on elderly people’s sense of safety and security. *Journal of Housing for the Elderly*, 20(3):45–60, 2006.
10. Sami Haddadin, Alin Albu-Schäffer, and Gerd Hirzinger. Safe physical human-robot interaction: measurements, analysis and new insights. In *Robotics research*, pages 395–407. Springer, 2010.
11. Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics*, 2(4):361–375, 2010.
12. Hiroko Kamide, Yasushi Mae, Koji Kawabe, Satoshi Shigemi, Masato Hirose, and Tatsuo Arai. New measurement of psychological safety for humanoid. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 49–56. ACM, 2012.
13. Andrew J Kornecki and Mingye Liu. Fault tree analysis for safety/security verification in aviation software. *Electronics*, 2(1):41–56, 2013.
14. Przemyslaw A Lasota, Gregory F Rossano, and Julie A Shah. Toward safe close-proximity human-robot interaction with standard industrial robots. In *Automation Science and Engineering (CASE), 2014 IEEE International Conference on*, pages 339–344. IEEE, 2014.
15. Christina Lichtenthäler, Tamara Lorenzy, and Alexandra Kirsch. Influence of legibility on perceived safety in a virtual human-robot path crossing task. In *RO-MAN, 2012 IEEE*, pages 676–681. IEEE, 2012.
16. A.H. Maslow. A theory of human motivation. *Psychological Review*, 50(4):370–396, 1943.
17. Seri Nonaka, Kenji Inoue, Tatsuo Arai, and Yasushi Mae. Evaluation of human sense of security for coexisting robots using virtual reality. 1st report: evaluation of pick and place motion of humanoid robots. In *Robotics and Automation, 2004. Proceedings. ICRA’04. 2004 IEEE International Conference on*, volume 3, pages 2770–2775. IEEE, 2004.
18. Ludovic Piètre-Cambacédès and Claude Chaudet. The SEMA referential framework: Avoiding ambiguities in the terms security and safety. *International Journal of Critical Infrastructure Protection*, 3(2):55–66, 2010.
19. James Russell. A circumplex model of affect. 39:1161–1178, 12 1980.
20. Theodoros Theodoridis and Huosheng Hu. Toward intelligent security robots: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):1219–1230, 2012.
21. Astrid Weiss and Christoph Bartneck. Meta analysis of the usage of the godspeed questionnaire series. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*, pages 381–388. IEEE, 2015.
22. K. A. Wyrobek, E. H. Berger, H. F. M. Van der Loos, and J. K. Salisbury. Towards a personal robotics development platform: Rationale and design of an intrinsically safe personal robot. In *2008 IEEE International Conference on Robotics and Automation*, pages 2165–2170, May 2008.