Evaluation of Human Perceived Safety during HRC Task using Multiple Data Collection Methods

Melis Sahin*
Pittsford Sutherland High School
Pittsford, NY, 14534, USA
melisgaripsahin@gmail.com

Celal Savur*
Electrical and Microelectronic Engineering
Rochester Institute of Technology
Rochester, NY, 14623, USA
cs1323@rit.edu

2022 17th Annual System of Systems Engineering Conference (SOSE) | 978-1-6654-9623-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/SOSE55472.2022.9812693

***Abstract*—While human safety is always a concern in an environment with human-robot collaboration, this concern magnifies when it is the human-robot work-space that overlaps. This overlap creates potential for collision which would reduce the safety of the system. Fear of such a collision could reduce the productivity of the system. This apprehensiveness is referred to as the perceived safety of the robot by the human. Therefore, we designed a within-subject human-robot collaboration experiment where a human and a robot work together in an assembling task. In order to evaluate the perceived safety during this HRC task, we collected subjective data by means of a questionnaire through two methods: during and after trial. The collected data was analyzed using non-parametric methods and these statistical tests were conducted: Friedman and Wilcoxon. The most clear relationship was found when changing only *sensitivity* of the robot or all three behaviors of *velocity*, *trajectory*, and *sensitivity*. There is a positive moderate linear relationship between the average safety of the during trial data and the after trial data.**

***Index Terms*—HRC, Data Collection Methods, Perceived Safety, Sensitivity**

I. Introduction

Recent years have seen a dramatic increase in situations where robots and humans must collaborate and co-exist [1]. In each of these situations, the biggest concern is safety. The priority of the robot is to ensure that the human does not get hurt. A situation in which direct contact occurs between a person and a robot is referred to as physical robot-human interaction (pHRI) [2], and is a great example of a system of systems. Not only should the robot not hurt the human, but the human should be confident that they are safe. Perceived safety describes the user's perception of the level of danger when interacting with a robot [3]. Achieving a positive perception of safety is a key requirement if robots are to be accepted as partners and co-workers [3]. If humans feel unsafe, whether they are actually in danger or not, they will likely be less productive. Therefore, it is imperative that we ensure robots are perceived as safe before we introduce and integrate them into our workplaces.

Perceived safety has generally been measured indirectly - most commonly by measuring physiological signals or through questionnaires. Unfortunately, there is not a unified method of measuring perceived safety across the field. Bartneck et al. [3] identifies numerous papers within the field which examined perceived safety that all used differing metrics. Most of the papers utilizing questionnaires collected the data post-trial. This is a weakness and limitation of the questionnaire method, as it allows for the distortion of the true feelings of participants due to errors in recollection which could bias their response [3].

In this research, in order to measure the perceived safety during HRC and to compare two data collection methods, we conducted an experiment with Sawyer, a collaborative robot. During the experiment, robot behavior was changed by manipulating the *velocity*, *trajectory*, and *sensitivity* parameters of the robot. Data was collected both throughout the trials as well as a single recorded response after the completion of the trial. The efficacy of the two data collection methods was analyzed.

The contribution of this paper is as follows:

- Evaluation of the perceived safety during HRC task
- Efficacy of two data collection methods
- Determine safer robot behaviors

The rest of the paper is structured as follows, section II provides a brief introduction to related works, section III introduces the methodology, and section IV defines the experiment. Section V provides and discusses results, and section VI discusses our conclusions.

II. Literature Review

Once it became feasible for robots to be integrated into society, the number one concern became safety. The primary focus of the field was developing safety measures, protocols, and features. The American National Standard for Industrial Robots and Robot Systems broadly categorizes the approaches to ensuring safety as reducing hazards through mechanical redesign, controlling the hazard through electronic safeguards, and warning the operator/user [4]. Safety in human-robot collaboration has been wildly studied, and much of what the field has to offer is collected in these survey papers [5], [6]. Collaborative interaction between humans and robots has two sides. Not only does the robot have to actually be safe, but in order for the person to work effectively, they must perceive the robot as safe.

Perceived safety is an aspect brought into academic focus relatively recently [7]. As such, there is not a unilaterally accepted methodology for evaluating human-robot interaction

* authors contributed equally to this research article.

978-1-6654-9623-0/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.

(HRI) studies. Commonly used measures are self-report, behavioral, psycho-physiological, and task performance-based. Self-report measures are the most frequently utilized method of evaluation in HRI studies, followed by physiological signals [8]. While questionnaires are the typical collection method for self-reported data, a key limitation identified in Bartneck et al. [3] is that the questionnaire is usually administered only after the actual experience. This means that subjects have to reflect on their experience afterward, which might bias their response. Recognizing this limitation, Zoghbi et al. [9] investigated the use of a hand-held device, coined the Affective-State Reporting Device (ASRD), that allows people to report their affective state, i.e., valence and arousal, continuously during a human-robot interaction trial. Their paper found that speed was the only significant factor on the subject's reported average arousal. Similarly, Koay et al. [10] attempted to directly measure a subject's comfort level via a simple device where subjects use a continuous scale to judge their current comfort level throughout an HRI interaction trial.

As a consequence of the lack of unified methodology, Bartneck et al. [3] found that researchers would use questionnaires with completely different scales, particularly Likert scales and semantic differential scales. Because of this, two studies that collected data using different scales cannot be accurately compared. In an attempt to remedy this, Navarro et al. [2] created OpenPHRI which is a C++/Python general-purpose software scheme with several built-in safety measures designed to ease robot programming for physical human-robot interaction and collaboration. Another unconventional method for comparing differing questionnaire scales was used by Kulic and Croft [11], [12]. Initially, they combined a questionnaire with physiological sensors to estimate the user's level of anxiety and surprise during sample interactions with an industrial robot. Data collected using a 5-point Likert scale displayed a strong positive correlation between anxiety and speed, and surprise and speed, and a negative correlation between calm and speed [11]. Later, they transformed their Likert template to the following semantic differential scales: Anxious/Relaxed, Agitated/Calm, Quiescent/Surprised and utilized with a new set of subjects to demonstrate that affective state arousal can be detected using physiological signals. In this way, Kulic and Croft were able to compare the results from two different questionnaires.

Nonaka et al. [13] describes a set of experiments where human response to pick-and-place motions of a virtual humanoid robot is evaluated. In their experiment, a virtual reality display is used to depict the robot. Human response is measured through heart rate measurements and subjective responses. No relationship was found between the heart rate and robot motion, but a correlation was reported between the robot velocity and the subject's rating of "fear" and "surprise" [13].

Overall, the most prevalent methodology is the combination of self-report measures and physiological signals. It is also important to remember that individual differences can create huge differences in perceived safety, as it is largely psychological. Since HRC includes two parties (i.e., humans and robots), safety perception is never based on the robot properties alone. Akalin et al. [14] showed that individual human characteristics, such as gender and personality traits, influence the perceived safety of humans in HRI. According to the data compiled in the survey paper by Rubagotti et al. [7], in general, for industrial manipulators, the feeling of perceived safety was enhanced when the relative human–robot distance was large, when the robot speed was low, and when controlling forces/avoiding abrupt robot motions during planned contact with the human.

In all the aforementioned methods and corresponding studies, the researchers used either after-trial or during-trial responses. In this study, not only did we measure perceived safety in a HRC task through a post-trial evaluation, but also during the trials. Doing this allows us to define safer robot behavior and determine efficient data collection methods.

III. Methodology

In this research, we investigate the effect of the robot's behavior on human perceived safety. In order to measure this effect, we programmed the robot to perform various behaviors, and asked the participants to fill out a subjective questionnaire on safety. The *velocity*, *trajectory* and *sensitivity* of the robot were manipulated while the participant interacted with the robot as shown in Fig. 1. Based on the results of the participant's self-reported questionnaire, the effect of the robot's behavior can be evaluated.

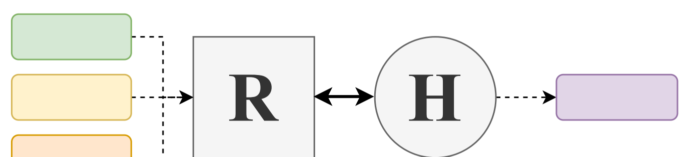

Fig. 1: Proposed work that investigate how changes in robot behavior affects human perceived safety.

One way to assess perceived safety is based on data reported by the human. Knowing how safe the person feels is critical for a HRC system and it can be used to program the robot's future behavior [15].

A. Robot Behaviors (Independent Variables)

During the experiment, robot behavior was altered by changing the *velocity*, *trajectory*, and *sensitivity* parameters of Sawyer. Velocity refers to the speed of the end-of-arm and has two levels: Normal (N) and Fast (F). Trajectory refers to the path of the arm and has two levels: Normal (N) and Extreme (E). Sensitivity refers to the hand-off and has two levels: Normal (N) and Sensitive (S). In addition, there is a random setting where a robot switches the *velocity*, *trajectory*, and *sensitivity* constantly and at-random during the duration of the movement. The different settings for each trial were denoted by a three letter arrangement. For each participant, we ran eight trials from Tab. I in addition to three RRR trials as shown in Tab. I.

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.

TABLE I: The list of the trials that participants involved during the experiment.

	Trial Type	# repeat	Description
1	*NNN*	1	Velocity *'normal'*, Trajectory *'normal'*, and Sensitivity *'normal'*
2	*NNS*	1	Velocity *'normal'*, Trajectory *'normal'*, and Sensitivity *'sensitive'*
3	*NEN*	1	Velocity *'normal'*, Trajectory *'extreme'*, and Sensitivity *'normal'*
4	*NES*	1	Velocity *'normal'*, Trajectory *'extreme'*, and Sensitivity *'sensitive'*
5	*FNN*	1	Velocity *'fast'*, Trajectory *'normal'*, and Sensitivity *'normal'*
6	*FNS*	1	Velocity *'fast'*, Trajectory *'normal'*, and Sensitivity *'sensitive'*
7	*FEN*	1	Velocity *'fast'*, Trajectory *'extreme'*, and Sensitivity *'normal'*
8	*FES*	1	Velocity *'fast'*, Trajectory *'extreme'*, and Sensitivity *'sensitive'*
9	*RRR*	3	Randomly selected from [1, 8] trial types.

B. Subjective Data Collection Methods

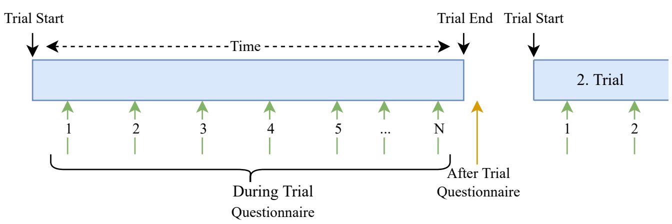

Fig. 2: Two data collection methods: During and After trial.

Fig. 2 shows the two data collection methods that were evaluated in this study. The first method is *after trial* wherein participants reported their perceived safety of the robot only after the conclusion of the trial. This method is very common in the field. The second method is *during trial* where the participants reported their perceived safety of the robot multiple times throughout the whole experiment.

In order to minimize the time that the participants spent reporting the subjective safety metric, a custom Android app was developed as shown in Fig. 3. The app allows the participant to enter their subjective responses immediately after the assembly of each part (iteration) with a single tab on the tablet screen. This minimizes the duration of the collection, maintaining the integrity of the experiment. The during trial approach produces more data and will provide a better idea of how the perceived safety is changing during the trials.

Fig. 3: The custom android app that participant enters the subjective metric by touching the bars on the screen.

C. Evaluation Criteria

IV. Experiment

In order to quantify perceived safety in human-robot collaboration, we conducted a sequential collaboration experiment [16] where the participant waits for the robot to bring a part for the assembly. The Sawyer [17] collaborative robot was used in the experiment and the data was collected from healthy college students (N=20). The participants consisted of 17 male and 3 female subjects (*Mean Age*= 24.70, *SD*= 2.99).

The experiment setup consisted of a joint task between the robot and the human, where the robot provides a part from Table-1 and the human picks a part from Table-2 as shown in Fig. 4. The participant is responsible for picking and screwing the two parts together, while the robot is holding one of the parts.

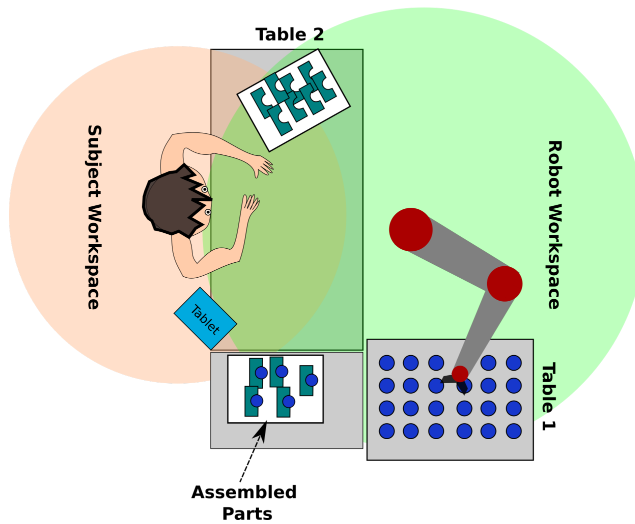

Fig. 4: Experiment 3: Stationary experiment, wait for the robot to bring the part.

In this experiment, we control the robot's behavior by changing its *velocity*, *trajectory*, and *sensitivity* (independent variables). The velocity was set to be in two levels of *normal* and *fast*. These two modes were achieved by setting the global speed ratio to 0.7 for normal mode and to 1.0 for fast mode. The trajectory was defined as two modes: *normal* and *extreme*. The normal trajectory was defined as the robot moving from Table-1 to Table-2 with minimum joint movements. Hence, the normal trajectory was smooth and predictable. On the other hand, the extreme trajectory passes multiple way-points between the pick and place locations which makes the robot jerkier and less predictable. The sensitivity was defined in two modes: *normal* and *sensitive*. The sensitivity was measured as the threshold of pressure on the robot when force was applied to the end-effector. Sensitive means there was a low threshold ($\pm 8N/s^2$) of pressure needed for the robot to move on to the drop location. Normal sensitivity means there was a high threshold ($\pm 11N/s^2$) of pressure needed for the robot to move on to the drop location.

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.

A. Experiment Procedure

This experiment consisted of 11 trials and one baseline. In a typical trial, the robot can provide up to 24 parts (iterations) to the participant. The robot job is to pick a part from *Table 1* and move it in front of the participant for assembling. The participant needs to pick an item from Table 2 and assemble the part in five seconds. Then, the robot moves to the assembled part location, as shown in 4, drops the part, and moves to Table 1 for the next part. If the participant is not able to assemble the part within five seconds, the robot will proceed to the dropping location. While the robot is dropping the part, the participant enters their responses to the tablet asking the subject to report *safety* levels in a continuous value range [0, 1] on the tablet screen. The trial is complete when five minutes have elapsed or there is no item left on Table 1 for the robot to pick up.

The participant was informed that they could tap the robot's end-effector to notify the robot not to wait anymore during assembling time. In this way, the participant could minimize the trial duration. Before the experiment started, all the participants were trained for one trial so that they got used to the task and the tablet. This minimized the impact that habituation and human acclimation may have had on the results of the experiment.

In addition to the subjective response during the trials, the participant filled a questionnaire after the completion of each trial where they reported their safety level, along with some other metrics.

V. Result and Discussion

In this section, we analyze multiple statistical tests to determine the effect of robot behavior on the perceived safety, and to compare two data collection methods.

A. During Trial vs. After Trial responses

Before starting any statistical analysis, we looked at information from a single trial. As we discussed in section III-B, the subjective metrics were collected during and after trials. The Fig. 5 shows the response of one of the participants from the study. The x-axis shows the number of the times the participant report the safety level and the y-axis shows the perceived safety level. The dashed line is the after trial response, and is the only value represented as a line.

In the Fig. 5 the during trial subjective response provides additional information about the safety of the participant. On the other hand, the after trial method summarizes the entire trial with only one reported value. Knowing this is critical to select which methods to use for a particular application. For example, for a physiological computing system that estimates safety based on physiological signals, the during trial data collection method would be a better approach [18]. On the other hand, for an application that evaluates overall robot behavior, it may be sufficient to use the after trial collection method.

The Pearson correlation between the average safety of during trial and after trial responses is 0.69. Hence, there is a moderate positive linear correlation existing between the two methods.

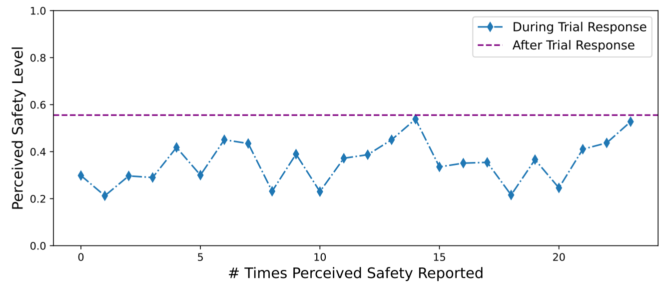

Fig. 5: A trial response of during and after trial from a random participant

Next, we will conduct multiple statistical analysis to investigate trial types.

B. Analysis of variance (ANOVA)

The ANOVA test is a commonly used method for statistical analysis when there are more than two groups. However, the underlying assumption for ANOVA test is that the values are normally distributed. We applied Shapiro-Wilk normality test to the perceived safety and all the trial types failed [19]. Hence, instead of ANOVA test we applied a non-parametric Friedman test which does not make any assumption about the underlying distribution for the reported values.

In this research, we used the test to determine whether there was a difference between trial types for both during and after trial response. The Friedman test results show that the $p\_val = 0.000169$, $p\_val = 0.000321$ for during trial and after trial respectively. This means that at least one trial type is different from the others. In order to know which trial type is different, a post-hoc test is needed.

C. Post-hoc Test

The post-hoc test provides one-to-one comparison between groups. We conducted a non-parametric paired t-test that uses Wilcoxon signed-rank test [20].

H_0: There is no difference in perceived safety between two trial types. For the sake of brevity, we will not repeat the null hypothesis for future comparison.

H_1: Participants would feel less safe when any of the *velocity*, *trajectory*, or *sensitivity* is different from *normal* $(H_1 : M_{NNN} > M_{FNN}, M_{NNN} > M_{NEN}, and M_{NNN} > M_{NNS})$.

As shown in Table II, we reject the null hypothesis (H_0) for the *NNS* since the *p_val* is less than $\alpha = 0.05$ for both methods. However, for the *FNN*, and *NEN*, while we fail to reject the null hypothesis for during trial, we reject the null hypothesis for after trial.

All other things kept normal (*NNN*), *sensitivity* was the only variable which had a significant effect on perceived safety based on during trial data. The other behaviors of *velocity* and *trajectory* only had a significant effect based on after trial

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.

TABLE II: Right-Tailed non-parametric paired Wilcoxon test between *NNN* and *NNS*, *FNN*, and *NEN*, respectively for both methods.

#	A	B	Cond.	During Trial (*p_val*)	After Trial (*p_val*)
1	NNN	NNS	*greater*	0.033	0.009
2	NNN	FNN	*greater*	0.415	0.001
3	NNN	NEN	*greater*	0.616	0.011

data. However, since there are limited data points, the after trial data collection method may be skewed. When adjusting robot behavior, *sensitivity* is most vital.

H_2: Participants would feel less safe when changing *trajectory* or *sensitivity* while velocity is fixed to *fast* (H_2 : $M_{FNN} > M_{FNS}, M_{FNN} > M_{FEN}$).

TABLE III: Right-Tailed Non-parametric paired Wilcoxon test between *FNN* and *FNS* and *FEN*, respectively for both methods.

#	A	B	Cond.	During Trial (*p_val*)	After Trial (*p_val*)
1	FNN	FNS	*greater*	0.000	0.029
2	FNN	FEN	*greater*	0.151	0.136

In this analysis, we kept the *velocity* fixed to *fast* and observed the impact of the *trajectory* and *sensitivity*. As shown in Tab. III, the test results showed that *FNS* is less safe based on both during and after trial responses. On the other hand, we failed to reject the null hypothesis for *FNN* and *FES* for both collection methods. The *p_val* is greater than the $\alpha = 0.05$. Similar to the previous analysis, we can see that sensitivity plays a vital role when *velocity* is *fast* as well.

H_3: Participants would feel less safe when changing *velocity* or *sensitivity* while trajectory is fixed to *extreme* (H_3 : $M_{NEN} > M_{FEN}, M_{NEN} > M_{NES}$).

TABLE IV: Right-Tailed non-parametric paired t-test between *NEN* and *FEN* and *NES*, respectively for both methods.

#	A	B	Cond.	During Trial (*p_val*)	After Trial (*p_val*)
1	NEN	FEN	*greater*	0.007	0.064
2	NEN	NES	*greater*	0.010	0.219

Similar to H_2, we kept the *trajectory extreme* and manipulate the *velocity* and *sensitivity*. As shown in Tab. IV, in this analysis, *FEN* and *NES* statistically significant which it means they are less safe than *NEN* on for during trial. Thus, we accept the alternative hypothesis. For the after trial we fail to reject H_0 due to limited evidence. The *p_vals* are low for the after trial this again may happen due to limited sample size.

H_4: Participants would feel less safe when changing *velocity* or *trajectory* while the sensitivity is fixed to the *extreme* (H_4 : $M_{NNS} > M_{FNS}, M_{NNS} > M_{NES}$).

As shown in Tab. V, the only test that was statistically significant was when $M_{NNS} > M_{FNS}$ for the after trial

TABLE V: Right-Tailed non-parametric paired t-test between *NNS* and *FNS* and *NES*, respectively for both methods.

#	A	B	Cond.	During Trial (*p_val*)	After Trial (*p_val*)
1	NNS	FNS	*greater*	0.059	0.035
2	NNS	NES	*greater*	0.335	0.219

response. The rest of the tests' *p_vals* were greater than $\alpha = 0.05$, hence we failed to reject the null hypothesis.

In this test, we set all independent variables to the extreme levels. The *velocity* was set to *fast*, *trajectory* was set to *extreme*, and the *sensitivity* was set to *sensitive*. We compared this robot behavior with other behaviors where we kept one variable fixed and changed the other two behaviors.

$H_5 : M_{FES} < M_{NES}, M_{FES} < M_{FNS}, M_{FES} < M_{FEN}$

TABLE VI: Left-Tailed non-parametric paired t-test between *FES* and *NES*, *FNS*, and *FEN*, respectively for both methods.

#	A	B	Cond.	During Trial (*p_val*)	After Trial (*p_val*)
1	FES	NES	*less*	0.002	0.041
2	FES	FNS	*less*	0.567	0.479
3	FES	FEN	*less*	0.059	0.057

Table VI shows the test results of *FES* vs. *NES*, *FNS*, and *FEN*. The table shows that *FES* was less safe than *NES* but there is not enough evidence to reject the null hypothesis for *FNS* and *FEN*. However, the *p_val* for $M_{FES} < M_{FEN}$ is very close to the $\alpha = 0.05$ value.

Based on all the statistical tests we conducted, FES is the least safe robot behavior. However, we conducted a trial type of *RRR* where the robot changes its behavior during the trial randomly from 1 to 8 in Tab. I.

H_6: Participants would feel least safe when robot behavior is *RRR* in comparison to all other trial types ($H_6 : M_{RRR} < M_{ALL}$).

The *p_vals* were 0.174 and 0.125 for during and after trial respectively. Hence, we failed to reject the null hypothesis. In order words, the *RRR* is not the least safe robot behavior.

VI. Conclusion

This research investigates the effect of the robot's behavior on the participants' perceived safety during human robot collaboration. An experiment was conducted involving HRC in order to accomplish a task. Data was collected from twenty participants to evaluate their perceived safety during and after the experiment. The pros and cons of the two methods were discussed.

The result of the experiment showed that robot behavior has an effect on the perceived safety. In addition, the Pearson correlation showed that there is moderate positive linear correlation between during and after trial response. From the subjective responses, we can see that each parameter of *velocity*, *trajectory*, and *sensitivity* of the robot has a different effect. From these three parameters, when we change only one

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.

parameter, the *sensitivity* is the most critical one that affects the perceived safety, followed by *trajectory*, and finally *velocity*. This paper found that for improved perceived safety, the better robot behavior recommended is normal sensitivity.

We also compared *RRR* trials, where the robot randomly chose from the independent variables, with the rest of the behaviors. There was not enough evidence to say that *RRR* was the least safe robot behavior.

In conclusion, when evaluating the perceived safety of a robot, during trial data will provide more information about the effect of the robot behavior on human perceived safety level. Alternatively, after trial data collection provides a sufficient overall perceived safety level. Based on the desired application, either of these methods can be useful. This paper compares two data collection methods to determine which one provides more accurate results.

Acknowledgment

The authors would like to thank Abhiraj Patil for his help during data collection. The authors are grateful to the staff of Multi-Agent Bio-Robotics Laboratory (MABL), the CM Collaborative Robotics (CMCR) Lab, and the Electrical Engineering Department at RIT for their valuable inputs. This material is based upon work partially supported by the National Science Foundation under Award No. DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] S. Kumar, C. Savur, and F. Sahin, “Survey of Human-Robot Collaboration in Industrial Settings: Awareness, Intelligence, and Compliance,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 280–297, 2021.

[2] B. Navarro, A. Fonte, P. Fraisse, G. Poisson, and A. Cherubini, “In Pursuit of Safety: An Open-Source Library for Physical Human-Robot Interaction,” *IEEE Robotics and Automation Magazine*, vol. 25, no. 2, pp. 39–50, 2018.

[3] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots,” *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[4] ANSI, “Industrial Robots And Robot Systems - Safety Requirements,” Tech. Rep., 1999.

[5] A. Pervez and J. Ryu, “Safe physical human robot interaction-past, present and future,” *Journal of Mechanical Science and Technology*, vol. 22, no. 3, pp. 469–483, 2008.

[6] S. Pinheiro, A. Correia Simões, A. Pinto, B. B. Van Acker, K. Bombeke, D. Romero, M. Vaz, and J. Santos, “Ergonomics and Safety in the Design of Industrial Collaborative Robotics,” in *Studies in Systems, Decision and Control*, ser. Studies in Systems, Decision and Control, P. M. Arezes, J. S. Baptista, P. Carneiro, J. Castelo Branco, N. Costa, J. Duarte, J. C. Guedes, R. B. Melo, A. S. Miguel, and G. Perestrelo, Eds. Cham: Springer International Publishing, 2022, vol. 406, pp. 465–478.

[7] M. Rubagotti, I. Tusseyeva, S. Baltabayeva, D. Summers, and A. Sandygulova, “Perceived Safety in Physical Human Robot Interaction – A Survey,” *Robotics and Autonomous Systems*, vol. 151, p. 104047, 2021. [Online]. Available: http://arxiv.org/abs/2105.14499

[8] C. L. Bethel, K. Salomon, R. R. Murphy, and J. L. Burke, “Survey of psychophysiology measurements applied to human-robot interaction,” *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 732–737, 2007.

[9] S. Zoghbi, D. Kulić, E. Croft, and M. Van Der Loos, “Evaluation of affective state estimations using an on-line reporting device during human-robot interactions,” *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 3742–3749, 2009.

[10] K. L. Koay, M. L. Walters, and K. Dautenhahn, “Methodological issues using a Comfort Level Device in human-robot interactions,” *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, vol. 2005, no. 1, pp. 359–364, 2005.

[11] D. Kulic and E. Croft, “Anxiety detection during human-robot interaction,” *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pp. 616–621, 2005.

[12] D. Kulić and E. Croft, “Physiological and subjective responses to articulated robot motion,” *Robotica*, vol. 25, no. 1, pp. 13–27, 2007.

[13] K. Inoue, S. Nonaka, Y. Ujiie, T. Takubo, and T. Arai, “Evaluation of human sense of security for coexisting robots using virtual reality (4th report: Comparison between real and virtual robots),” *Proceedings of the SICE Annual Conference*, pp. 3879–3882, 2005.

[14] N. Akalin, A. Kristoffersson, and A. Loutfi, “Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures,” *International Journal of Human Computer Studies*, vol. 158, no. November 2021, p. 102744, 2022. [Online]. Available: https://doi.org/10.1016/j.ijhcs.2021.102744

[15] N. Sarkar, “Psychophysiological control architecture for human-robot coordination - Concepts and initial experiments,” *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 4, no. May, pp. 3719–3724, 2002.

[16] A. BAUER, D. WOLLHERR, and M. BUSS, “HUMAN–ROBOT COLLABORATION: A SURVEY,” *International Journal of Humanoid Robotics*, vol. 05, no. 01, pp. 47–66, mar 2008.

[17] Hann, “Sawyer Robot,” 2018. [Online]. Available: https://www.rethinkrobotics.com/sawyer/

[18] C. Savur, S. Kumar, and F. Sahin, “A framework for monitoring human physiological response during human robot collaborative task,” *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, vol. 2019-Octob, pp. 385–390, 2019.

[19] S. S. Shapiro and M. B. Wilk, “An Analysis of Variance Test for Normality (Complete Samples),” *Biometrika*, vol. 52, no. 3/4, p. 591, dec 1965. [Online]. Available: https://www.jstor.org/stable/2333709?origin=crossref

[20] F. Wilcoxon, “Individual Comparisons by Ranking Methods,” *Biometrics Bulletin*, vol. 1, no. 6, p. 80, dec 1945. [Online]. Available: https://www.jstor.org/stable/10.2307/3001968?origin=crossref

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on July 19,2022 at 13:37:06 UTC from IEEE Xplore. Restrictions apply.