Applying the Perceived Creepiness of Technology Scale to Social Robots

Jessica Turner
University of Waikato
Tauranga, New Zealand
jessica.turner@waikato.ac.nz

Judy Bowen
University of Waikato
Hamilton, New Zealand
judy.bowen@waikato.ac.nz

Jemma König
University of Waikato
Hamilton, New Zealand
jemma.konig@waikato.ac.nz

Katarzyna Stawarz
Cardiff University
Cardiff, United Kingdom
StawarzK@cardiff.ac.uk

Nicholas Vanderschantz
University of Waikato
Hamilton, New Zealand
vtwoz@waikato.ac.nz

***Abstract*—Designing positive robot experiences requires an understanding of users' perceptions and meeting their needs in an ethical manner. However, despite best intentions, users have strong positive or negative reactions to robots, either finding them "cute" or "creepy". The Perceived Creepiness of Technology Scale (PCTS) was designed for evaluating how creepy a technology appears to a user on first encounter. In this paper we applied the PCTS to a cross-section of social robots to measure their perceived creepiness and evaluate the strengths and weaknesses of PCTS when applied in a Human-Robot Interaction (HRI) context. We demonstrate that while a robot may not be perceived as creepy initially, it can have underlying unethical practices inherent in its design which is not well captured by the PCTS. This emphasises the need for better HRI practices to ensure creepiness is appropriately assessed in the social robot domain.**

***Index Terms*—creepiness; creepy; perceived creepiness technology scale; social robots; human robot interaction.**

I. Introduction

From service robots used in restaurants, to companion robots used to combat loneliness, social robots are fast appearing in our everyday lives [1]. As such, it is necessary to ensure that the robots we introduce into human spaces are ethical, enable positive user experiences, and meet users' needs. End users tend to have strong reactions to robots, either finding them "cute" or "creepy", depending on their appearance, behaviours, and interactions [2]. While there are many factors that make technology creepy, such as invasion of privacy or concerns around ethical boundaries [3], social robots present new challenges as they can move and act autonomously within our environment.

Whether users find technology creepy or cute has a direct impact on their perceptions and experience. In [4] the authors describe how users' perceptions of social robots are shaped by society, culture and popular media. Tools like the Godspeed questionnaire by Bartneck et al. [5] allow us to measure and understand user perceptions of robotics to determine the impact of design choices. Conversely, the Perceived Creepiness Technology Scale (PCTS) was designed for evaluating how creepy a technology appears on first encounter and to enable discussion amongst designers to determine creepiness levels and design better, less creepy systems [6].

To our knowledge, there are no scales or frameworks which evaluate perceived creepiness in Human-Robot Interaction (HRI). While Godspeed provides general understanding of robot perceptions, it does not evaluate creepiness specifically. Conversely, PCTS aims to generalise creepiness to cover all technologies and does not consider the unique challenges introduced by interactions with social robots. In this paper we explore the use of PCTS to evaluate the perceived creepiness of social robots and the associated challenges with defining what makes a "creepy robot". The primary contribution of this paper is an analysis of PCTS which provides insights for an adapted perceived creepiness scale in the HRI domain. In particular, our results demonstrate that key aspects related to creepiness or cuteness of social robots, including anthropomorphism, appearance, terminology, data privacy and ethics, are not reflected well in the PCTS scale, emphasising the need for new HRI techniques.

In this paper we start by describing related work and techniques for investigating robot user perceptions. Next we describe selection of several social robots which are analysed using the PCTS. We then provide recommendations for creepy robot analysis and outline the need for further techniques to compliment existing tools.

II. Related Work

A. Human-Robot Interaction and Social Robots

The advances in HRI and increasing prevalence of robots, especially in settings like restaurants, hotels, or hospitals [7]–[9], call for the development of new tools and techniques to design novel user experiences between robots and humans. Bartneck et al. [10] give an overview of HRI techniques, challenges, and state of current research in the field highlighting the complexities around design and evaluation. Furthermore, HRI research categorises robots based on their use in society. Guizzo et al. [11] define 18 categories, including educational robots, medical robots, and social robots. Our focus within this paper is on social robots. As defined by the authors in [12], social robots “emphasize social interaction with users as their main affordance”. Their primary function is to socialise with other robots and/or humans in a variety of settings.

Mahdi et al. [13] conducted a systematic review of literature, with the purpose of surveying trends in social robot design. As part of their work, they define five categories of social robots. In our work, we evaluate four socially situated robots, two robots that operate within an environment they can perceive and respond to, and two socially interactive robots, where social interaction is central to their purpose. In addition, Mahdi et al. define 6 key domains of application: service, healthcare, education, research, entertainment and telepresence. We use these domains of application for robot selection in Section III.

B. *“Creepy” Technology*

As technology integrates more closely into our lives and homes, researchers have recognised the risk of introducing feelings of unease or discomfort. In [6] the authors refer to this as “creepy” technology. Researchers have investigated the creepiness of well-being apps [14], artificial intelligence with human-like personas [15], and children’s perception of creepy technology [2], to name a few. Kang et al. investigate user perception of “creepy” versus “cool” service robots [16]. Within their work, they evaluate three categories of service robot: machine-like, semi-human-like, and human-like. While they did not illustrate a strong difference in perception of creepiness between the three categories, they did highlight an association between perception and value. They found that perceived creepiness was negatively linked with perceived value, and that perceived coolness was positively linked with perceived value. That is, the creepier the robot was seen to be, the less useful users felt it was. Similarly, Brink et al. discuss the “uncanny valley” effect [17]. The term, originally coined by Mori, suggests that a robot’s creepiness increases when it resembles a human in appearance, speech, and behavior [18]. The uncanny valley metaphor is well accepted in adult perceptions and Brink et al. have found the same is true for children [17]. They suggest that uncanny valley develops over time and is linked to evolving perceptions of robot consciousness [17]. This highlights the important link between appearance and perceived creepiness of robots and the impact it has on user experience.

C. *Techniques for HRI Evaluation*

There are several tools used to evaluate robots and related user perceptions. One of the most popular is Godspeed developed by Bartneck et al. [5], which allows researchers to compare and validate results [19] to understand differences in robot perceptions. In contrast, in [20] the authors present a framework for evaluating the functionality of robots to understand which to incorporate in different scenarios of use. Given the importance of trust in HRI [21], [22], Saari et al. devised a robot-focused version of the Technology Acceptance Model questionnaire that prioritises factors related to perceived usefulness, perceived ease of use, robot’s self-efficacy, and user’s robot anxiety, among others [23]. Similarly, in [24] Schaefer develops the Trust Perception Scale (TPS) to be used in HRI, while in contrast Vaiani et al. give an overview on end-user development in HRI and approaches to ensure non-expert users are involved in the development process so that they meet users needs [25]. While each of these tools allows us to explore different aspects of HRI and robot design, we could not find any that focus on assessing creepiness specifically in the HRI domain. Given the strong link between perceived creepiness of robots and their perceived value, it is evident that analysis techniques for creepiness in HRI will be key to the success of social robotics moving forward. Therefore, our aim was to evaluate whether PCTS [6], the existing creepiness scale, could also be applied to social robots.

TABLE I: The Perceived Technology Creepiness Scale [6]

Subscale/Item
Implied Malice, α = 0.83
Q1: I think that the designer of this system had immoral intentions.
Q2: The design of this system is unethical.
Undesirability, α = 0.75
Q3: Using this system in public areas will make other people laugh at me.
Q4: I would feel uneasy wearing this system in public.
Q5: The system looks bizarre to me.
Unpredictability, α = 0.80
Q6: This system looks as expected. (R)
Q7: I don’t know what the purpose of the system is.
Q8: This system has a clear purpose. (R)

III. Methodology

The PCTS defined in [6] is used to understand how creepy technology is perceived to be when first encountered. The scale builds on a conceptual model by using best practice identified in literature and qualitative data gathered from focus groups. The full 8-item scale is described in Table I. The scale is intended to be used by designers, researchers, or user study participants to rate each of the items from strongly agree (7) to strongly disagree (1). The lowest possible score on the scale is 9 (least creepy) with the highest score being 63 (most creepy).

To evaluate the effectiveness of the scale, we selected 6 commercially available robots and used the PCTS to evaluate their perceived creepiness (see Fig. 1). For each robot, we collected a name, photo, webpage and advertising video; all were used as part of the evaluation. The authors of the paper served as the five evaluators for this study. Each evaluator completed the PCTS for each robot in a randomised order to prevent bias in the results. During this process the evaluators also took notes and identified any issues they encountered while performing their evaluation. At the end of the study, each evaluator rated the robots from most to least creepy without reference to the final calculated scores. The objective of this analysis was to determine the applicability of PCTS to social robots and whether or not it allowed for fair comparison.

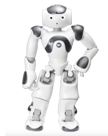
During robot selection we prioritised robots with different functionalities to get a representative dataset. Based on the key domains of application of social robots identified by Mahdi et al. (service, healthcare, education, research, entertainment,

(a) AV 1

(b) Bittle

(c) ElliQ

(d) KettyBot



(e) Nao

(f) Pepper

Fig. 1: Social Robots used in Study

telepresence) [13], we identified robots from each domain (see Fig. 1). **AV 1** is a telepresence robot used in classrooms to allow children unable to attend school to still be active in the learning process. **Bittle** is an educational robot designed as a companion where the user can teach it new tricks and functionalities as their ability in programming improves. **ElliQ** is a companion robot targeted at the aging population to help reduce isolation and improve wellbeing. **KettyBot** is a service robot that can perform the roles of an usher, waiter, marketer, or tour guide, and assist humans in commercial settings. **Nao** is a fully programmable humanoid robot widely used in research to support social interactions [26]. Similarly, **Pepper** is also a humanoid service robot that users can interact with via voice and touch screen. In contrast to Nao, Pepper is much larger and also has a screen for displaying relevant information.

IV. Analysis and Results

Based on the evaluation, the KettyBot was considered the least creepy with ElliQ being the most creepy, although no scores were above 40, indicating that the robots were not considered creepy. The range of scores for each robot varied by a maximum of 14.5 to a minimum of 7, indicating the level of agreement between the evaluators. The results for each robot are described below.

ElliQ had the highest creepiness scores (average=32.5) and range of 11.5 points between the scores, indicating some minor disagreement between the evaluators. They highlighted that the nature of interaction was important when determining creepiness. As ElliQ collects personal data, this created a feeling of uncertainty around where that data was going and concerns around invasion of privacy. The use of vague language such as "promotes health" without concrete evidence of how the system achieves this also made it harder to trust that the data is protected and used appropriately. This was especially concerning given the robot is targeted towards aging adults. ElliQ triggered feelings of discomfort around the reliance on robots to provide companionship. The evaluators also stated that they would feel uncomfortable using the robot in public. Furthermore, the gender assigned to the robot ("she") raised concerns given its subservient role.

Bittle, who has a clearly defined purpose as an educational device, was generally perceived as cute by the majority of the evaluators (average=27.4), with 12 points between the highest and lowest scores, indicating some disagreement between the evaluators. When watching the Bittle videos, multiple evaluators noted shots that included a headless version of the robot, which impacted their answers. Similarly, the combination of a semi-realistic head with a robot body seemed out of place. Two evaluators also noted a likeness to similar robots used by the Police or Army, which contributed to its perceived creepiness. For one of them the overall small size and customisability limited this impact, while for the other, the realistic face of the dog (which appeared to scowl or growl) heightened the sense of unease.

Nao received an average score of 26.7 with a range of 14.5 points between the highest and lowest scores, emphasising a significant division between evaluators. This placed it as 3rd least creepiest overall. The evaluators who found Nao creepy noted that robots designed as human replacements are more creepy than those with well defined roles or tasks. Again, issues around face and head design appeared, with one evaluator finding Nao's red and green eyes off-putting as they appeared to be always open and never blinking. Those who found Nao less creepy did so due to its cuteness and size, comparing it to a small toy. Prior exposure also mattered: evaluators mentioned that seeing the robot on social media prior to the study reduced the perceived level of creepiness. One evaluator also mentioned Nao and Pepper for research, which contributed to reduced PCTS scores for both robots.

Pepper received an average score of 26.5, with a range of 12.5 points between the highest and lowest scores, indicating some difference of opinion between the evaluators. Pepper's website shows it hugging a child, and this slow awkward movement was found to be "the creepiest thing I've seen any of the robots do!" by one of the evaluators. The size of the robot was also important for the evaluators who compared it to a small child. Similar to Nao, the robot's eyes appear to be always open, although the video of Pepper clearly depicted it "blinking" by dimming the lighting of the eyes momentarily, which helped to address some of the creepiness found in Nao.

AV 1 was the 2nd least creepy robot (average score = 26.2), with 8 points difference between the highest and lowest one, indicating a high level of agreement between the evaluators. The evaluators noted that the 'privacy first' model of AV 1 and its focus on users' needs directly contributed to it being less creepy. One evaluator thought that the animation used in the video was less creepy than just a picture of the robot itself, which had an impact on their perceptions. However, since the robot does not show any likeness to the child it is representing, there were concerns that it could lead to bullying, as other children in the classroom may forget they are interacting with a real person, which could lead to further isolation.

KettyBot received an average score of 25, with a range of 7 points between the highest and lowest scores, indicating general agreement between the evaluators. However, all evaluators had used or interacted with a KettyBot before. One evaluator noted that the robot's less humanoid design makes it less creepy, while in contrast others found heads and faces important to reduce creepiness. The KettyBot materials showed two face options: one with a single eye and another with a full face. One evaluator found the single eye design creepier, which changed their perception of the robot, highlighting the impact of faces on creepiness.

Each evaluator was asked to separately rank the robots from most to least creepy and their rankings were compared with the PCTS scores. For 4 evaluators there were significant differences between their self-ranking and the ranking determined by PCTS. None of the evaluators' rankings were fully aligned, with the most creepy robot identified as ElliQ, Pepper, AV1 or Bittle. Similarly, KettyBot, AV1, Bittle or Nao were seen as the least creepy based on self rankings. Note that AV1 and Bittle appear in both the most and least creepy robots rankings, further suggesting that perceived creepiness is subjective.

V. Discussion

The primary aim of the PCTS is to allow evaluators to determine the levels of creepiness of a particular technology [6]. The application of the PCTS to robots has provided a clear tool for comparison between these devices; however, it has also highlighted aspects of creepiness not considered by the scale due to the unique characteristics of social robotics.

A. Cuteness, Creepiness & Privacy

The PCTS highlights implied malice, undesirability and unpredictability as the key aspects for creepiness [6]. However, our results demonstrate that functionality is also important. Evaluators noted that robots with a clearly defined function were perceived as less creepy than those with unclear functionality. This is reflected in the results, with ElliQ and Bittle receiving higher PCTS scores than AV 1 and KettyBot, which have a defined purpose and where the user has less freedom to control their behaviour.

Anthropomorphism is directly linked with the emotional responses that arise when people nurture, care for or engage with robots [13]. All evaluators contrasted their notes about "creepiness" with either "cuteness" or "coolness", and these comments related directly to anthropomorphism. Appearance and design issues relating to head and faces were linked to this "cuteness" factor, along with size. It is evident from this analysis that the robot's head, face, size and design are important alongside clearly defined functionality.

Lastly, data privacy and protection were noted by all evaluators in relation to ethical concerns, which is a recognised concern in the context of social robots [27]. AV 1's use of data protection contributed to its lower PCTS scores, while ElliQ's unclear data protection had an opposite effect. The function and appearance elements described here are not considered directly in the PCTS, although can be seen as part of the undesirability and unpredictability aspects of the scale.

B. Evaluating Social Robots with PCTS

All evaluators noted that being exposed to a robot can have an impact on its perceived creepiness. We believe this is reflected in lower scores for robots the evaluators had interacted with before (as evidenced by KettyBot's scores). The order in which the material for each robot was reviewed also had an impact on the results.

As the PCTS was designed to consider technology more generally and be applicable to a wide range of systems, the terminology did not always apply to social robots. For example, Q4 of the scale, "I would feel uneasy wearing this system in public", did not make sense in this context as none of the robots evaluated could be physically worn by the user. Each evaluator translated this in a slightly different way when using the scale, from taking the question literally to assuming this meant interacting with the robot in public.

However, the PCTS did enable discussions amongst the evaluators as intended with relation to creepiness, highlighting the need for a similar scale in the HRI domain. Specifically, the results demonstrate the importance of elements used in the Godspeed questionnaire [5] such as anthropomorphism and likeability alongside those used in the PCTS, like implied malice, undesirability and unpredictability. Beyond updating terminology to ensure the scale can be applied in the HRI domain, we recommend the addition of aesthetics to ensure the robot is appealing and adds to its cuteness and likeability factors. Furthermore, how the data was shared and used was also important in the creepiness factor. To reduce creepiness, developers must follow best ethical practices and provide users with control and freedom over their own data. Ethical design is also key to ensure that the robot meets the users' needs and has a clearly defined purpose.

VI. Conclusion and Future Work

In this paper we evaluated a representative sample of social robots using the PCTS to determine its applicability to the HRI domain. Evaluators applied the scale to each robot and the relevant PCTS scores were calculated alongside individual rankings of the robots by each evaluator from most to least creepy. We found that PCTS allowed us to reason about the creepiness level of each robot but that the scale rankings did not match well with users rankings. There were issues in PCTS with regards to anthropomorphism, appearance, terminology, data privacy and ethics being key aspects of creepiness for social robots that were not reflected well in the scale. However, this evaluation highlighted the need for tools like the PCTS to evaluate creepiness of social robots in the HRI domain. This paper provides clear motivation for further development of a PCTS scale for social robots in combination with popular tools like the Godspeed questionnaire. These techniques would allow for an appropriate framework for evaluating robot creepiness, understanding user perceptions, designing usable robots and ensuring robots meet users' needs.

References

[1] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *International Journal of Social Robotics*, vol. 5, pp. 291–308, 2013.

[2] J. C. Yip, K. Sobel, X. Gao, A. M. Hishikawa, A. Lim, L. Meng, R. F. Ofiana, J. Park, and A. Hiniker, "Laughing is scary, but farting is cute: A conceptual model of children's perspectives of creepy technologies," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–15.

[3] R. R. McWhorter and E. E. Bennett, "Creepy technologies and the privacy issues of invasive technologies," in *Research anthology on privatizing and securing data*. IGI Global, 2021, pp. 1726–1745.

[4] E. Sobolewska, N. Vanderschantz, and F. Bitcon, ""he's adorable and i want to take him home". trust perceptions before and after first-time encounters with social robots," 2021.

[5] C. Bartneck, "Godspeed questionnaire series: Translations and usage," in *International handbook of behavioral health assessment*. Springer, 2023, pp. 1–35.

[6] P. W. Woźniak, J. Karolus, F. Lang, C. Eckerth, J. Schöning, Y. Rogers, and J. Niess, "Creepy technology: what is it and how do you measure it?" in *Proceedings of the 2021 CHI conference on human factors in computing systems*, 2021, pp. 1–13.

[7] C. A. Cifuentes, M. J. Pinto, N. Céspedes, and M. Múnera, "Social robots in therapy and care," *Current Robotics Reports*, vol. 1, pp. 59–74, 2020.

[8] L. Ragno, A. Borboni, F. Vannetti, C. Amici, and N. Cusano, "Application of social robots in healthcare: Review on characteristics, requirements, technical solutions," *Sensors*, vol. 23, no. 15, p. 6820, 2023.

[9] A. Rosete, B. Soares, J. Salvadorinho, J. Reis, and M. Amorim, "Service robots in the hospitality industry: An exploratory literature review," in *Exploring Service Science: 10th International Conference, IESS 2020, Porto, Portugal, February 5–7, 2020, Proceedings 10*. Springer, 2020, pp. 174–186.

[10] C. Bartneck, T. Belpaeme, F. Eyssel, T. Kanda, M. Keijsers, and S. Šabanović, *Human-robot interaction: An introduction*. Cambridge University Press, 2020.

[11] E. Guizzo, "Types of robots: Categories frequently used to classify robots," May 2023. [Online]. Available: https://robotsguide.com/learn/types-of-robots

[12] J. Favela, D. Cruz-Sandoval, M. Marques da Rocha, and D. C. Muchaluat-Saade, "Social robots for healthcare and education in latin america," *Communications of the ACM*, 2024.

[13] H. Mahdi, S. A. Akgun, S. Saleh, and K. Dautenhahn, "A survey on the design and evolution of social robots—past, present and future," *Robotics and Autonomous Systems*, vol. 156, p. 104193, 2022.

[14] J. S. Seberger, H. Choung, J. Snyder, and P. David, "Better living through creepy technology? exploring tensions between a novel class of well-being apps and affective discomfort in app culture," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, no. CSCW1, pp. 1–39, 2024.

[15] R. R. McWhorter and E. E. Bennett, "Creepy technologies and the privacy issues of invasive technologies," in *Research anthology on privatizing and securing data*. IGI Global, 2021, pp. 1726–1745.

[16] S.-E. Kang, C. Koo, and N. Chung, "Creepy vs. cool: Switching from human staff to service robots in the hospitality industry," *International Journal of Hospitality Management*, vol. 111, p. 103479, 2023.

[17] K. A. Brink, K. Gray, and H. M. Wellman, "Creepiness creeps in: Uncanny valley feelings are acquired in childhood," *Child development*, vol. 90, no. 4, pp. 1202–1214, 2019.

[18] M. Mori, "Bukimi no tani [the uncanny valley]." *Energy*, vol. 7, p. 33, 1970.

[19] A. Weiss and C. Bartneck, "Meta analysis of the usage of the godspeed questionnaire series," in *2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 2015, pp. 381–388.

[20] L. Onnasch and E. Roesler, "A taxonomy to structure and analyze human–robot interaction," *International Journal of Social Robotics*, vol. 13, no. 4, pp. 833–849, 2021.

[21] B. C. Kok and H. Soh, "Trust in robots: Challenges and opportunities," *Current Robotics Reports*, vol. 1, no. 4, pp. 297–309, 2020.

[22] D. R. Billings, K. E. Schaefer, J. Y. Chen, and P. A. Hancock, "Human-robot interaction: developing trust in robots," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 2012, pp. 109–110.

[23] U. A. Saari, A. Tossavainen, K. Kaipainen, and S. J. Mäkinen, "Exploring factors influencing the acceptance of social robots among early adopters and mass market representatives," *Robotics and Autonomous Systems*, vol. 151, p. 104033, 2022.

[24] K. E. Schaefer, "Measuring trust in human robot interactions: Development of the "trust perception scale-hri"," in *Robust intelligence and trust in autonomous systems*. Springer, 2016, pp. 191–218.

[25] G. Vaiani and F. Paternò, "End-user development for human-robot interaction: Results and trends in an emerging field," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, no. EICS, pp. 1–40, 2024.

[26] A. Robaczewski, J. Bouchard, K. Bouchard, and S. Gaboury, "Socially assistive robots: The specific case of the nao," *International Journal of Social Robotics*, vol. 13, pp. 795–831, 2021.

[27] E. Fosch-Villaronga, C. Lutz, and A. Tamò-Larrieux, "Gathering expert opinions for social robots' ethical, legal, and societal concerns: Findings from four international workshops," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 441–458, 2020.