

A Case Study of Paro: Addressing the Trade-off Between Transparency and Well-Being

GROUP 33 MEMBERS:

AMIN BERJAOUI TAHMAZ

MSc Robotics, Robot & Society, 5610737

EDOARDO PANICHI

MSc Robotics, Robot & Society, 5630444

GIOVANNI CORVI

MSc Robotics, Robot & Society, 5608899

AND

GIOELE BURIANI

MSc Robotics, Robot & Society, 5629888

[Revised on 11 April 2022]

1. Introduction

Some of the recently developed controversial technologies are in the field of robotics, especially in the domain of socially assistive robots. Paro, developed by the Advanced Industrial Science and Technology (AIST), is an interactive, fur-covered robot that resembles a baby harp seal. It falls into the category of therapeutic robots, also called carebots, and it was designed to provide companionship in nursing home settings. According to Paro's official website [1], its main characteristic is that it appears responsive to its environment and the interactions it receives from it, which proved to be extremely successful with dementia patients. Their interactions with it are similar to when they play with pets: they might pet it, talk to it or even sing to it, which has a calming effect on the patients. In particular, nurses noted that interactions with Paro encouraged social engagement while reducing stress and negative emotions, resulting in a better-perceived care experience. What's more, considering the growing market of socially assistive robots, it becomes clear how this technology could increase access to care services on a global scale.

However, it is worth noting that Paro's introduction to the care system has not been without controversy. Many people saw it as a powerful tool of emotional seduction that could deceive mentally fragile patients. For them, unfortunately, the line between imagination and reality is often blurry, which might inevitably make this technology excessively powerful and misleading. While the potential of PARO and other carebots is recognized by everybody, some notable critics, including the virtue ethicist and philosophy professor Shannon Valor, point out the potential risks involved with robot care if applied without first going through an ethical examination process [2].

Another focal point is related to whether the patients should be warned about the true nature of their companion. Kathy Craig, a therapist at the Livermore Veteran's Hospital where Paro has been used for almost 10 years, said that she does not explicitly tell the patients Paro the seal is a robot, stating that "we find it works better with people with dementia because if the residents are aware that it's not real, we find that sometimes they don't engage with it as much" [2]. This obviously raises even more

concerns about the ethical aspect of the implementation of this robot. This paper aims to tackle one of the controversies surrounding Paro. More specifically, we aim to answer: ‘Do the benefits of socially assistive robots (Paro) outweigh the ethical concerns associated with developing an illusory relationship with dementia patients?’ or, more specifically, ‘Is it ethically admissible to sacrifice Paro’s transparency for the patient’s well-being?’

In the following sections, we will first give a brief literature review regarding the current literature on the topic, as well as an overview of two ethical frameworks that could be implemented to try and analyze this situation: value sensitive design and sociotechnical systems analysis. After that, we will analyze the aspects of transparency and trust on an individual patient level, studying whether there is a value breach due to the introduction of Paro. Lastly, we will see the implications of Paro on a societal level and the potential value shift that may take place due to its deployment on a national scale.

2. Analysis and Evaluation

In order to tackle this research question, we intend to take on two frameworks: value sensitive design and sociotechnical systems analysis.

Value sensitive design can be defined as ‘a way of doing ethics that aims at making moral values part of technological design, research and development’ [3]. In other words, it is a framework which incentivizes the embedding of ethical values within the design of a technology. In particular, the technology that is being discussed in this paper deals with some extremely delicate values in our society, such as mental illness and healthcare in general. Therefore, we think that VSD could give some particularly interesting insights since it ‘offers a coherent method for evaluating the current design of technologies’ [4]. Also, this technology is still at an early stage of development and distribution, so VSD could also help shaping it in the right direction as it ‘offers a proactive element to influence the design of technologies early on and throughout the design and implementation process’ [4].

On the other hand, sociotechnical systems analysis deals with analysing the problem from the perspective of a sociotechnical system, defined as an integral intertwinement of the technical, social and cultural components that mutually inform and co-produce one another. This analysis is clearly applicable to the case of Paro (technology) since it is designed to have a strong interaction with dementia patients (social actors) under the supervision of hospitals and nursing homes (institutions). Our analysis will try and determine whether this interaction is beneficial for all the parts, and therefore should be encouraged, or if it needs to be modified due to some ethical concerns.

As for existing literature, there is an abundance of research on socially assistive robots in elderly care [5][6][7]. Also, there exists some papers looking into value conflicts regarding transparency and utility [8] and their implications on a societal level [9][10]. However, a gap exists when tackling the value tradeoff in Paro’s particular case. This is an especially important matter since countries such as the US and Japan cover Paro as a therapy device under their health insurance policies. All in all, we can make use of the existing literature to gain insight and make meaningful inferences regarding the step forward in either promoting or discouraging Paro’s use in elderly homes and healthcare facilities.

In the following sections, we will attempt to answer the aforementioned research question. This will be done by first analyzing Paro using a sociotechnical approach by examining Paro’s effect on three stakeholders: the technology itself, the dementia patients, and the elderly community as a whole. By examining each stakeholder, we can identify the ethical implications of Paro’s lack of transparency on a focused scale, allowing us to further investigate any potential misuse. For the first stakeholder (technology), we discuss the technology behind Paro and highlight how it leads to a breach in transparency. For the second stakeholder (dementia patients), we discuss the patients’ state

with and without introducing Paro, and elaborate on which value should be prioritized given Paro's untransparency. Lastly, we highlight the impact of Paro on the third stakeholder (elderly community) in order to infer the implications of this impact.

2.0.1. First Stakeholder - The Technology, Paro:

According to the official website, Paro is an interactive robot used as a therapeutic device for dementia patients [1]. It is designed to look like a baby harp seal in order to incorporate the benefits of animal therapy, without the need to introduce live animals into a care facility.

Moreover, Paro makes use of five sensors that allow it to detect light, interaction forces, sounds, temperature, and even the posture of its handler. This allows it to respond to user interactions, such as calling its name or petting it, which helps develop a personal connection when interacting with it. However, Paro's presentation as a live, interactive animal brings about a potential ethical problem: untransparency.

Before making any claims on whether Paro exhibits transparency, we must first define it. According to Wortham et al., transparency entails the extent to which a robot's capacities, constraints, and intentions are comprehended by the users [11]. This understanding can go beyond physical actions and extend into the robot's capacity to reciprocate care. In the case of Paro, studies showed that its simulated sentience is perceived as real by the dementia patients, with patients describing Paro as a 'buddy' and 'good companion' [12]. One might make the argument that untransparency is not unethical due to the absence of ill-intention, the latter of which is attributed to deceit [13]. However, this does not negate the existence of false beliefs embedded as a design parameter according to VSD, since the false perception regarding Paro's sentience is still present in the minds of the patients. Taking it a step further, Sparrow argues that self-deception stems from developing an illusory relationship with a robot conflicts with the need to experience the actual world, implying that there exists an evident transparency breach when such a relationship develops [14]. Thus, Paro's design is made with disregard to the value of transparency from a VSD perspective.

2.0.2. Second Stakeholder - Dementia Patients:

In order to adequately analyze the dementia patients as a stakeholder, we must first establish a medical definition of dementia: Dementia stems from a loss in cognitive ability, mainly in the form of thinking, recalling, and reasoning [15]. This loss is classified as dementia when it affects the person's independence and ability to do everyday routines. In turn, this leads to several symptoms that include, but are not limited to: memory loss, confusion, paranoia, and viewing everyday things as unfamiliar [15].

The article further elaborates that the stakeholders (dementia patients) become mentally frail and require constant attention from family members or caregivers. As mentioned in the introduction, patients interacting with Paro experience a reduction in stress and an increase in overall psychological well-being. Moreover, it further encourages the patients to interact with their caretakers while simultaneously improving their social interaction within the other patients at the care facility [1].

While Paro's introduction into care facilities brings about several benefits, it may be in conflict with the values of the stakeholders. In the previous section, we established that Paro is untransparent. However, this disvalue is an unintended consequence of the desire to promote the patient's well-being. According to value sensitive design, this value conflict between transparency and health falls into the category of moral overload, in which a compromise of one value is necessary in order to promote another value [3]. Kuran et al. elaborates on a set of useful strategies to tackle moral overload, and

the most relevant strategy to this case is casuistry [16]. Casuistry entails that there exists special circumstances under which a value being sacrificed for the sake of another value can be exempted.

In the remainder of the section, we will look into which value should be sacrificed and the reasoning behind it. More specifically, we intend to associate primary values with the concepts of transparency and well-being, and accordingly rank these values along a hierarchy. In turn, this will provide us insight into whether transparency is worth being compromised for the sake of the patient's well-being.

In the context of elderly care, Grodzinsky et al. state that transparency in robots most notably appears as the value of trust [11]. In turn, a robot breaches the value by acting as a live animal since it deceives them into believing that it is indeed a live animal. On the other side of the spectrum, well-being can be effectively modeled using the value of health [17]. Since the two values are competitively entangled, one must take their vulnerable mental and physical state into consideration, as well as the potential worsening of their condition unless appropriate treatment is provided. Thus, the admittance of a dementia patient into a healthcare facility necessitates that the patient's health is prioritized over other values along the hierarchy, including trust.

However, a breach of trust must only be permissible as long as this breach falls within an established set of constraints. Yew et al. suggest a set of relevant constraints applied to carebots, which can be used in our context: the intention of the robot developer aims to enhance the patient's health, the end result is a net positive impact on the patient, the deception embodied by the robot must be necessary, and there exists no other alternatives within the robot itself that allow it to achieve the same result [11]. Looking into Paro, it is evident that it abides by the four constraints, and is consequently ethical in its behavior.

Pairing the fragile state of the dementia patients with our conclusion of prioritizing health over transparency, we find it permissible for Paro to compromise transparency in favour of the patients' well-being.

2.0.3. Third Stakeholder - Elderly Community:

In this paper, we consider the elderly community as a whole to be the third stakeholder. Contrary to the two previous sections, we will first identify the main advantages and disadvantages of Paro's use, then generalize the mentioned points to assess its impact on the aforementioned stakeholders.

When Paro is used on a large scale, then its individual benefits and consequences would extend beyond the singular level and into a societal level. Assuming Paro becomes a common therapeutic device, one can infer the potential shift in values on the current stakeholder level by looking into the implications of Paro's benefits and consequences.

After concluding that Paro's behavior is ethical, we find it preserves the values of individual dementia patients. However, when Paro is used collectively by a community of dementia patients, then its individual benefits and consequences would extend beyond the singular level and into a societal level. Assuming Paro becomes a common therapeutic device, one can infer the potential shift in values on the current stakeholder level by looking into the implications of those benefits and consequences.

Following an array of clinical trials and through closely following the use of Paro by dementia patients, the following key benefits were identified. Firstly, Paro is made to simulate the effect of animal therapy, as such the results consisted of reducing negative emotion and behavioral symptoms in dementia patients, as well as improving their overall social engagement. According to a study by Richeson et al., using Paro as a simulation of animal-assisted therapy resulted in a significant decrease in agitated behaviors [18]. Moreover, the improvement in mental health due to Paro's introduction can be seen to have had a tangible impact on those same patients, whereby there was a reduction in the prescribed pain and behavior medication [19].

Paro also paves the way for social justice, which is defined as the equal access to rights and opportunities within a society. This access should be granted to all the economic classes of society. Countries such as the US and Japan have labeled Paro as a therapeutic device which can be claimed under the national health insurance policy [20]. In turn, Paro represents a case study of the potential for socially assistive robots in terms of promoting the quality of life of dementia patients, regardless of their financial state.

However, Paro's introduction into elderly care as a common therapeutic device will also have two main consequences. Due to Paro's pacifying companionship, the human caregivers will naturally devote less attention to the patients with Paro as compared to patients that need their care. However, this reduction of human contact might lead to isolation and social complications in the long run [21]. Another interesting insight arose after reading surveys regarding experiences with Paro [22][23]. The participants stated that Paro is effectively a toy and found it demeaning to be treated like children. They further expressed feelings of embarrassment about interacting with Paro. Lastly, a minority of the elderly had frustrated and angry responses due to their bad past experience with animals.

After laying out the main advantages and disadvantages of introducing Paro into dementia care, we will now assess their implications in terms of societal values. The first implication is apparent, which is the promotion of the value of health for the elderly community by looking into engineering solutions to partially remedy their medical problems. Moreover, the notion of equality will be further propagated by endorsing and applying these applications that fall under social justice across other diseases prevalent in the elderly community. In terms of the impact of reduced human contact, this may lead to social complications between the dementia patients and their caretakers. Pairing this with the infantilizing of the patients, when placed on a societal level, it may lead to a decline of the elderly community's perceived dignity [24][25].

One might argue that these values will not be as pronounced since they are associated with a minority relative to the society. However, our response is that we make no claim that the introduction of Paro to this minority will shift societal values. These societal values are dynamic and a change in them is a result in multiple incremental value shifts, eventually resulting in a noticeable change. Thus, by using Paro as a case study, we can get insight into the values we might be propagating in order to maximize the beneficial ones while minimizing detrimental ones.

3. Conclusion

In this paper, we tried to analyze the possible ethical repercussions of technologies such as Paro, an interactive robot resembling a harp seal pup, used to treat elderly people with dementia. The identified ethical problem lies in the necessity of the deception in order for Paro to provide utility, such that the patients are fooled into believing that they are interacting with a real animal. As we have said, this is a clear violation of important values such as transparency and trust, but at the same time, we have shown that within the frameworks of value sensitive design and sociotechnical systems analysis, these can be sacrificed to promote a higher value: *health*.

This analysis may have been enough to unravel the ethical issues underlying Paro's use from three different stakeholder perspectives: the technology itself, a dementia patient, and the elderly community as a whole. We argued that Paro as a technology is untransparent, but that this untransparency should not compromise a more important value in this context, which is the patient's health. Furthermore, we reason that the benefits and consequences of Paro's introduction will have a ripple effect into the societal value system, which is the form of increased access to elderly care paired with a decline in their perceived dignity.

The procedure we have used to investigate PARO has to be intended as a generalizable approach for ethics in robotics and therefore can be implemented to solve similar problems and case studies that extend to the domain of socially assistive robots as a whole. This paves the way for a variety of new questions, most notably: what are the contexts in which well being can be sacrificed for transparency, if any? In addition, who can decide whether the patient being deceived is providing consent for this value compromise (for the sake of their well-being)?

We believe this work will enable researchers and practitioners to conduct thoughtful ethical analysis in a field as vibrant and active as robotics, where innovation is everyday life.

REFERENCES

1. PARORobots PARO Therapeutic Robot. (2014), <http://www.parorobots.com>
2. A.Johnson Robotic seals comfort dementia patients but raise ethical concerns. (2014), <https://www.kalw.org/show/crosscurrents/2015-08-17/robotic-seals-comfort-dementia-patients-but-raise-ethical-concerns>
3. Hoven, J., Lokhorst, G. & Poel, I. Engineering and the problem of moral overload. *Science And Engineering Ethics*. **18**, 143-155 (2012)
4. Van Wynsberghe, A. Designing robots for care: Care centered value-sensitive design. *Science And Engineering Ethics*. **19**, 407-433 (2013)
5. Calo, C., Hunt-Bull, N., Lewis, L. & Metzler, T. Ethical implications of using the paro robot, with a focus on dementia patient care. *Workshops At The Twenty-fifth AAAI Conference On Artificial Intelligence*. (2011)
6. Johnston, A. Robotic seals comfort dementia patients but raise ethical concerns. *KALW Local Public Radio*. (2015)
7. Vandemeulebroucke, T., Casterlé, B., Welbergen, L., Massart, M. & Gastmans, C. The ethics of socially assistive robots in aged care. A focus group study with older adults in Flanders, Belgium. *The Journals Of Gerontology: Series B*. **75**, 1996-2007 (2020)
8. Ishowo-Oloko, F., Bonnefon, J., Soroye, Z., Crandall, J., Rahwan, I. & Rahwan, T. Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation. *Nature Machine Intelligence*. **1**, 517-521 (2019)
9. Gerłowska, J., Skrobas, U., Grabowska-Aleksandrowicz, Korchut, A., Szklener, S., Szcześniak-Stańczyk, D., Tzovaras, D., & Rejdak, K. Assessment of perceived attractiveness, usability, and societal impact of a multimodal robotic assistant for aging patients with memory impairments. *Frontiers In Neurology*. **9** pp. 392 (2018)
10. Langer, A., Feingold-Polak, R., Mueller, O., Kellmeyer, P. & Levy-Tzedek, S. Trust in socially assistive robots: Considerations for use in rehabilitation. *Neuroscience & Biobehavioral Reviews*. **104** pp. 231-239 (2019)
11. Yew, G. Trust in and ethical design of carebots: the case for ethics of care. *International Journal Of Social Robotics*. **13**, 629-645 (2021)
12. Hung, L., Gregorio, M., Mann, J., Wallsworth, C., Horne, N., Berndt, A., Liu, C., Woldum, E., Au-Yeung, A. & Chaudhury, H. Exploring the perceptions of people with dementia about the social robot PARO in a hospital setting. *Dementia*. **20**, 485-504 (2021)
13. Donath, J. Ethical issues in our relationship with artificial entities. (Oxford University Press Oxford, UK, 2020)
14. Sparrow, R. & Sparrow, L. In the hands of machines? The future of aged care. *Minds And Machines*. **16**, 141-161 (2006)
15. NIA What Is Dementia? Symptoms, Types, and Diagnosis. (2021), <https://www.nia.nih.gov/health/what-is-dementia>
16. Kuran, T. Moral overload and its alleviation. *Economics, Values, And Organization*. pp. 231-266 (1998)
17. Dobewall, H., Tark, R. & Aavik, T. Health as a value and its association with health-related quality of life, mental health, physical health, and subjective well-being. *Applied Research In Quality Of Life*. **13**, 859-872

- (2018)
18. Richeson, N. Effects of animal-assisted therapy on agitated behaviors and social interactions of older adults with dementia. *American Journal Of Alzheimer's Disease & Other Dementias*®. **18**, 353-358 (2003)
 19. Petersen, S., Houston, S., Qin, H., Tague, C. & Studley, J. The utilization of robotic pets in dementia care. *Journal Of Alzheimer's Disease*. **55**, 569-574 (2017)
 20. Shibata, T. Life Innovation with Therapeutic Robot, Paro. (2012)
 21. Tergesen, A. & Inada, M. It's not a stuffed animal, it's a \$6,000 medical device. *The Wall Street Journal*. **6**, 1 (2010)
 22. Moyle, W., Bramble, M., Jones, C. & Murfield, J. Care staff perceptions of a social robot called Paro and a look-alike Plush Toy: a descriptive qualitative approach. *Aging & Mental Health*. **22**, 330-335 (2018)
 23. Moyle, W., Bramble, M., Jones, C. & Murfield, J. "She had a smile on her face as wide as the great Australian bite": a qualitative examination of family perceptions of a therapeutic robot and a plush toy. *The Gerontologist*. **59**, 177-185 (2019)
 24. Vercelli, A., Rainero, I., Ciferri, L., Boido, M. & Pirri, F. Robots in elderly care. *DigitCult-Scientific Journal On Digital Cultures*. **2**, 37-50 (2018)
 25. Sharkey, A. & Sharkey, N. Granny and the robots: ethical issues in robot care for the elderly. *Ethics And Information Technology*. **14**, 27-40 (2012)
 26. Ann Light and Rosemary Luckin. *Designing for social justice: people, technology, learning*. 2008.
 27. *What is Social Justice?* <https://corporatefinanceinstitute.com/resources/knowledge/other/social-justice/>
 28. Vijaya Murali and Femi Oyeboode. *Poverty, social inequality and mental health*. Advances in Psychiatric Treatment, 2004.
 29. *Basic psychologist, free service* <https://www.breakinglatest.news/health/basic-psychologist-free-service/>