Helpless by Design:
Helplessness as a Feature in Human-Robot Interaction

Maarten H. Lamers*
Peter van der Putten*
m.h.lamers@liacs.leidenuniv.nl
p.w.h.van.der.putten@liacs.leidenuniv.nl
[A]Social Creatures Lab, LIACS, Leiden University
Leiden, The Netherlands

ABSTRACT

In situations where robots fail at their normal functioning, they may become unable to operate further without help from others. However, helplessness need not be the results of failing only, and can be purposefully designed into robots as a feature rather than an undesired effect or consequence. We describe how helplessness-by-design was explored and exploited in a human-robot interaction (HRI) context, in both academia and artistic practice. Our overview aims to initiate discussion and provide a knowledge-base about robot helplessness in HRI.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design theory, concepts and paradigms**.

KEYWORDS

Robots, failure, helplessness, human-robot interaction

1 INTRODUCTION

Robots can obviously fail, but when does it become not a fault, but a key feature? When designing helpless robots a lack of competence, power and capability becomes the focus, not a side effect to account for.

Robots are generally envisioned as functional, goal-driven physical agents. Certainly, failure to achieve their task or operate within their intended requirements, is something that we can all imagine. And the more robots are employed in the wild, or in complex and varying environments, the more our imagining of robotic failure becomes expectation.

*Both authors contributed equally to this research.

Such failure may threaten the robot's operational existence. It may expose them to threatening conditions such as environmental factors (e.g. water, or lack of sunlight for solar powered operation), or violation of basic operating conditions (e.g. being unable to reach a recharging station). In such cases one could envision the robot being in a state of helplessness: being unable to guarantee their continued operational or intended existence without help from others.

In environments where humans and robots collocate, people may be confronted with robots in a state of helplessness while being able to help them overcome their adverse situation. Consequently, such scenarios may raise questions about people's ability and willingness to help a robot in need, or the risk that it may bring to them personally. In turn these questions may expose more fundamental issues, such as the projected value of robot "life" or operation, or the prevalence of negative bias towards robots. Or issues regarding ethics, such as whether it is ethical to not assist a helpless robot. As such we argue that helplessness is an aspect of robot behavior or operation that is relevant to the field of human-robot interaction (HRI).

Accepting robot helplessness as a topic deserving attention from the HRI community opens a realm of research questions. These broadly fall into the overlapping domains of human behavior understanding, robot design, and ethics. One could study factors underlying people's willingness to help a robot in need, such as robot appearance (e.g. zoomorphism vs abstraction), expressive motion [2], projected emotion (e.g. fear or anger), vocalization, collaboration (e.g. when the robot explains how it can be helped), or even bribery (e.g. offering reward in return for help). Also, employed research methods could be diverse. Human responses to helpless robots could be studied in lab settings, or in more ecologically valid in-the-wild environments. Methods could be exploratory or closed (e.g. hypothesis driven). In any case, there is plenty of opportunity and need to study robot helplessness in an HRI context.

In this paper we explore how helplessness was designed into robots purposefully, as a feature rather than an undesired effect or consequence. We focus on how this helplessness-by-design was explored and exploited in an HRI context. Our overview aims to open the discussion about robot helplessness in HRI work, and to provide an initial knowledge base about purposefully designed robot helplessness. We do not attempt to define what is a robot and are very accepting of different forms. For the purpose of our study, it is acceptable to include non-physical agents such as chat-bots, or artifacts that appear to people as robots, since the HRI reflection that they allow would likely generalize to physical agents also.

Figure 1: *The Helpless Robot* **(1987-2002) by Norman White (photo: Wikimedia Commons, CC BY-SA 3.0)**

Although our work is framed in an academic context, we draw from both academic and artistic practices. Artists serve humanity with alternative realities and critical reflections on any thinkable topic. Robots and HRI are fruitful subjects for artists. Not only have robots been used as artifacts to address societal issues via artistic works, but also critical reflections on HRI itself were implemented in the form of robots. And such works may serve us well besides academic output to uncover and understand robot helplessness.

This work is derived from a larger project by the authors that compiles a collection of examples of robots with qualities that are traditionally associated with humans, not robots.[1] This collection is divided into multiple categories, such as (but not limited to) misbehaving robots (e.g. racist, smoking), religious robots, defecating robots, and helpless robots. Naturally it is from the latter category that most works in this paper were drawn.

In the remainder of this short paper we present examples of helpless robots (Section 2) and discuss their implications along the way. We show how designed helplessness can hold value within the field of HRI. Section 3 concludes the paper.

2 HELPLESS ROBOTS

In this section we present examples of robots that were designed to be helpless, sometimes partial, sometimes in full. We call these robots helpless robots, but they were selected for having helplessness play a role in the relation between humans and robots. Our aim is not to be exhaustive, in that other purposefully helpless robots may exist, but to present an overview representative of different

[1]See https://www.botslikeyou.org

forms of and reasons for robot helplessness. Moreover, categorization is tentative and not mutually exclusive. It was applied mainly to structure the narrative.

2.1 Care and Nurturing

Popular visions of the future often incorporate robots designed to care for humans, such as nursing robot *Robear* (RIKEN and Sumitomo Riko Company, 2015) that cares for elderly with reduced mobility, and other rehabilitation robots that can assist skilled clinicians [1]. In such visions, the care is directed from the robot towards humans. But naturally, care and compassion can be directed from human towards a robot also. Designed robot helplessness can stimulate care-giving behaviour in people, which in turn could serve multiple human-oriented goals.

One such goal could be entertainment. *Tamagotchi* digital pet devices introduced by Bandai (1996) feature designed helplessness for human entertainment. A digital pet represented inside the device would go through several stages of development when properly cared for by a human caregiver. Without proper care, the creature could not survive and would ultimately "die". Caring in this context meant regular feeding, playing with the pet, disciplining, cleaning, and even medicating. From the enormous popularity of the devices, one may argue that the designed helplessness of Tamagotchi digital pets manages to interest and entertain people.

Another classic example is PARO, a cute baby seal that requires human care. It is a therapeutic robot that offers animal assisted therapy aiming to reduce patient stress and stimulate interaction and socialization. Target patients include those suffering Alzheimer's disease. Patients can pet PARO and talk to it, to which it responds as if it were a living baby seal, moving its head, limbs and body, making sounds, and showing what behavior it prefers from patients, imitating the vocality of a real baby harp seal.

Whilst there is debate about ethical aspects (e.g. [7]) and actual performance [3], PARO was clearly designed with helplessness as a feature in mind. The intention is not to replace human care, but to be an alternative for a pet or a cuddly toy. Pets are often not allowed in care homes, and cuddly toys may lack just enough of a trigger to activate patients. Patients may feel helpless themselves: they may experience a loss of agency, power and control, and are hitting boundaries with respect to their abilities. So rather than being cared for, it can be nice to turn the tables and have something that requires their help.

2.2 Power Dynamics

The robot with perhaps the most appropriate name in our overview is *The Helpless Robot* (various instantiations existed between 1987-2002) by media artist Norman White. Positioned in an artistic context, the robot was intended to interact with museum/exhibition visitors. With a decidedly non-zoomorphic human-sized appearance (Fig. 1), it quite resembled an abstract *Dalek* from BBC television series *Doctor Who*. Standing in a gallery environment, its electronically synthesized voice asked for the physical help of passers-by; with a friendly and persuasive tone it would request visitors to rotate it, a purpose for which it was equipped with handles. As such the robot gave the impression of helplessness to visitors: expressing a desire (to be rotated), but not the ability to do so by itself.

Figure 2: *Colored Sculpture* **(2016) by Jordan Wolfson, at Stedelijk Museum Amsterdam (photo: Peter van der Putten)**

Figure 3: *Can't Help Myself* **(2016) by Sun Yuan and Peng Yu (photo: Ana Romero López via Wikimedia Commons, CC BY-SA 4.0)**

While being helped by visitors, the voice of the robot slowly changed to a more forceful, commanding tone, complaining when the assistance was not being executed properly or as fast as desired. As such the work confronts visitors with a potential future in which robots not only help us, but may also demand our help. Moreover, it plays with the dynamics of power in the context of care-giving. As the care is provided, the robot attempts to distort/reverse the power-relation on which the care was initially based, thereby adding a dimension to the public's experience and interpretation of helplessness and the complex and potentially ambiguous nature of power relationships.

2.3 Powerless Captivity

A more intense version of helplessness is being powerless or lacking the ability to (re-)obtain required power to overcome a given adverse process or situation. And certainly powerlessness is caused by captivity. We identify multiple robotic works that play with the concept of powerless captivity.

One such robotic installation is the artwork *Colored Sculpture* (2016) by Jordan Wolfson. It consists of a larger-than-life cartoon-like puppet of a boy, chained at one hand, one foot and from the top of its head, being alternately suspended, dropped and dragged over the floor by an agent in the form a room-sized truss structure (Fig. 2). The positions in which the boy-puppet is put convey to the audience a feeling of helpless discomfort and powerlessness.

At first glance, robotic agency belongs only to the chain-controlling structure, with the boy-puppet a powerless captive victim. However, the puppet is equipped with electronic displays for eyes. Through facial recognition software they search out the public's faces and stare directly and intensely at them, with a menacing and unmoved gaze. This blurs feelings of helplessness ascribed by the public to the boy-puppet, and muddles the boundaries of agency and control between the structure and the puppet. Are they really one? And if so, what remains of the powerlessness projected onto the boy-puppet, as he literally keeps the audience captive? Are we, the public, being haplessly mislead? As with *The Helpless*
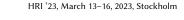
Robot, the power-relation between the robotic artwork and public is questioned and possibly reversed.

Powerless captivity is an element also projected by viewers onto the robotic artwork *Can't Help Myself* (2016) by artists Sun Yuan and Peng Yu. A large KUKA industrial robot caged inside a brightly lit transparent acrylic enclosure, stands in what appears as a pool of blood on the floor (Fig. 3). With a squeegee-like appendix to its arm, it attempts to sweep the dark red liquid into a perfect circle around itself; a truly Sisyphean task, as the liquid flows to break the circular shape. Reacting to visitors' presence, its powerless behavior is occasionally interrupted by a series of dance-like moves, giving it a conscious sense of its surroundings and a more human-relatable identity, until it continues its mindless, caged, and endless task.

Many possible interpretations of *Can't Help Myself* have been proposed by critics (discussed in [5]). In most of these, the work was clearly designed to provoke a sense of robot-helplessness in the viewer, at which the work's title also hints. When in November 2021 videos of the work went viral on TikTok, it touched a nerve in many netizens, who expressed their thoughts on social media. Many saw their own lives reflected in the work, and interpreted the robot's captive powerlessness in contexts such as addiction to social media and capitalist exploitation of workers [5].

Other artists have also presented robots as powerless "caged animals". In fact, media artist Bram Ellens created an exhibition of his robot works titled *Robots in Captivity* (2021). Varying types of robots were placed in various forms of captivity, to reflect on possible human-robot relationships.

2.4 Helpless In The Wild

In contrast to being in captivity, helplessness can also be the result of being "lost in the wild", perhaps with no specific aim, or with an aim but not the means to fulfill it. This section presents several robots that were purposefully built to roam helplessly in an uncontrolled environment cohabited by humans and dependent on human help through chance encounters. From an HRI research perspective, such methods could overcome experimental issues of ecological validity.
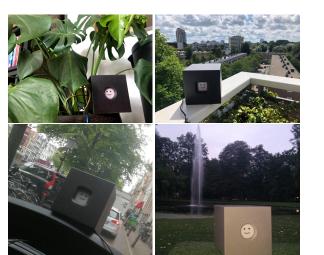
Figure 4: BlockBot on the road (photos: from [4] with permission)

In 2014 *hitchBOT* by David Harris Smith and Frauke Zeller successfully hitchhiked its way across Canada, and later also in other countries. As a combination of "research-creation" and performance art (as stated by its makers [8]) it was meant to stimulate reflection on trust and empathy between human and robot. In short, a technically unsophisticated immobile robot would depend on strangers offering it rides to reach an intended location. Helplessness was the result of it depending on human help to achieve its aims, and of its lacking protective systems such as a human chaperon. This aspect became painfully explicit when *hitchBOT* was found to be destroyed through vandalism on one of its travels in 2015.

Having a robot roam in the wild is a useful method to study HRI in its envisioned human ecology, as robot helplessness may serve as a driver of social interaction. Other works have taken a similar approach, in which a helpless robot "abandoned" in public space depends on the kindness of strangers to get to where it intends to go. Kacie Kinzer's *Tweenbots* (2009) are but one example.

This method was also employed in an HRI study that investigated the impact of a robot and humans sharing a common locus, hypothesized to strengthen the perception of shared experiences which in turn increases social bonding. Helpless cube-shaped 'couch-surfing' robotic artifacts with minimal human appearance or behavior (Fig. 4) were passed from host to host to gather information about common locus and human-robot social bonding [4]. The study found that people's bringing of the BlockBot into their personal home, workplace, etc, would correlate to more identity and mind attributions made to these minimal cube shaped artifacts.

3 DISCUSSION

Our paper does not present a crafted empirical study, but an initial exploration into how robot helplessness was exploited in prior studies and artistic practices. And clearly, our exploration is not exhaustive in that other works involving helpless robots may exist.

From the examples described we can assess that designed helplessness in robots holds value for HRI. It provides methods for engaging people with robots, for example through care and nurturing of robots, which can in turn serve multiple goals such as education, entertainment, and therapy. But also dealing with helpless robots can teach us about possible futures with robots in our daily lives. How would we respond to a failing robot, and what factors modulate that response? How willing are we to bring an unknown robot into our lives?

In fact, how robot helplessness relates to human/animal helplessness is an interesting research topic that was not covered in our overview. No formal research could yet be uncovered, although an interesting research was proposed in 1999 by Freedom Baird, student at the time [9]. She theorized that expressions of fear by *Furby* robotic toys (Tiger Electronics, 1998) when held helplessly upside down, would make people identify them partly as a creature, not only a machine, and act differently as a consequence of that despite being aware of its machine-nature. When the experiment was informally executed in 2011 as part of a *Radiolab* podcast episode [6], children were asked to hold upside down a Barbie doll, a Furby, and a real hamster, while measuring how long it took before their emotions made them want to turn it back upright. Naturally the Barbie doll does not complain, while Furby's whining and vocal expressions of fear when held upside down may be interpreted as indications of an inner life more alike the hamster than the Barbie. As hypothesized by Baird, children would upright the helpless Furby much faster than the Barbie, closer to how fast they would for the hamster, supporting her theory.

We reviewed works that purposefully designed helplessness into robots, as a feature rather than an undesired effect or consequence, plus we provide an initial knowledge base about purposefully designed robot helplessness. As such, we hope to open discussion and possibly further research, and conclude that robot helplessness is a potential factor to consider in HRI research.

REFERENCES

[1] Roberto Colombo and Vittorio Sanguineti. 2018. *Rehabilitation Robotics: Technology and Application*. Academic Press. https://doi.org/10.1016/C2016-0-02285-4

[2] Guy Hoffman and Wendy Ju. 2014. Designing Robots with Movement in Mind. *J. Hum.-Robot Interact.* 3, 1 (feb 2014), 91–122. https://doi.org/10.5898/JHRI.3.1.Hoffman

[3] Hee Sun Kang, Kiyoko Makimoto, Rie Konno, and In Soon Koh. 2020. Review of Outcome Measures in PARO Robot Intervention Studies for Dementia Care. *Geriatric Nursing* 41, 3 (2020), 207–214. https://doi.org/10.1016/j.gerinurse.2019.09.003

[4] Joost Mollen, Peter van der Putten, and Kate Darling. 2023. Bonding with a Couchsurfing Robot: The Impact of Common Locus on Human-Robot Bonding In-the-Wild. *J. Hum.-Robot Interact.* 12, 1, Article 8 (feb 2023), 33 pages. https://doi.org/10.1145/3563702

[5] Iris Olde Hampsink. 2022. Can't Help Myself – How a Relatable Robot Offers a Critical Reflection on Modern Society. *Diggit Magazine* (2022). https://www.diggitmagazine.com/papers/can-t-help-myself-how-relatable-robot-offers-critical-reflection-modern-society

[6] Radiolab Podcast. 2011. *Furbidden Knowledge*. Retrieved January 30, 2023 from https://radiolab.org/episodes/137469-furbidden-knowledge

[7] Amanda Sharkey and Natalie Wood. 2014. The PARO Seal Robot: Demeaning or Enabling?. In *Proceedings of AISB 2014* (London, UK). Curran Associates, Inc., 284–288.

[8] David Harris Smith and Frauke Zeller. 2017. The Death and Lives of hitchBOT: The Design and Implementation of a Hitchhiking Robot. *Leonardo* 50, 1 (02 2017), 77–78. https://doi.org/10.1162/LEON_a_01354

[9] Sherry Turkle. 2012. *Alone Together*. Basic Books, New York, NY.