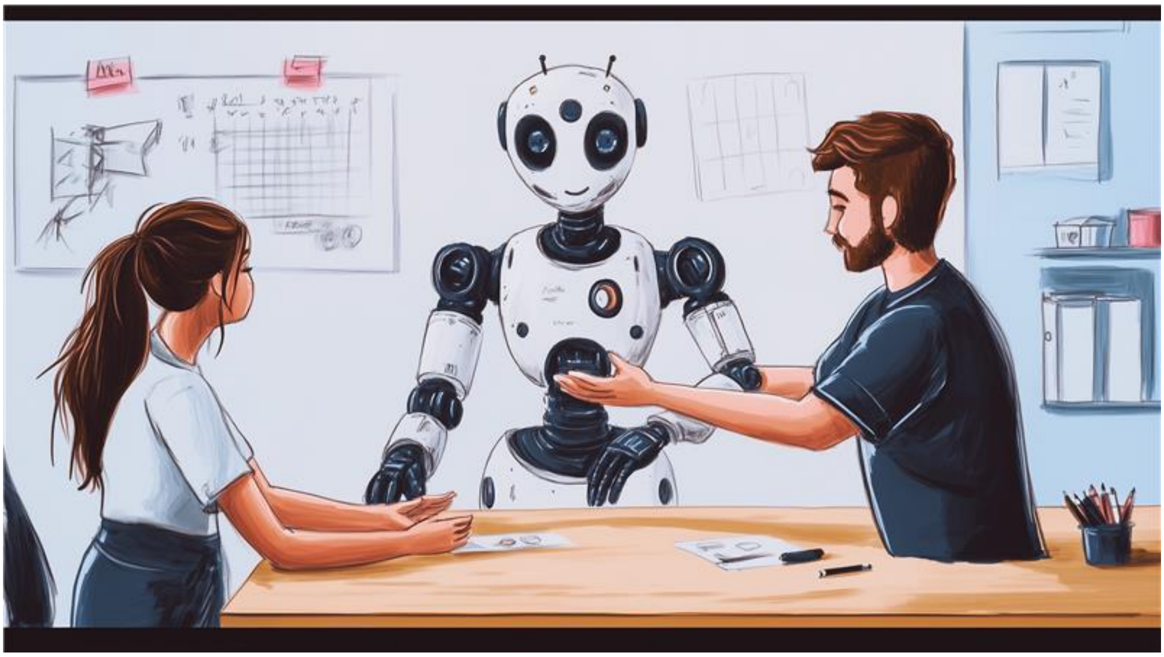
SURF

Driving innovation together

Social Robots in education

Should we prepare for social robots in education?



Author: Merel van Steenis (Intern Responsible Tech)
Supervisors: Duuk Baten and John Walker
Contact: responsibletech@surf.nl

Version: 1.0
Date: 03-03-2025
Identifier:

Table of contents

Introduction

This October 2024 at Tesla's 'We Robot' event in Hollywood, Elon Musk showcased the new advancements of the AI-powered humanoid robot Optimus[1]. A robot designed to perform a wide range of tasks in many different contexts. According to Musk, Optimus could be 'a teacher, babysitter, walk your dog, mow your lawn, get the groceries, serve you drinks, or whatever you can think of: Optimus will walk amongst people'[2]. This new technological innovation highlights the possibility that advanced humanoid robots, once maybe even mainstream, might reshape our ways of living tremendously.

As often with similar technologies, the jump to its potential in education is quickly made[3]. Educational influencers such as Dan Fitzpatrick suggest Optimus could revolutionize education as a teaching assistant by preparing materials, supervising students, and reducing the administrative burden on teachers[4]. Optimus' price tag of 20.000 to 30.000 dollars makes it plausible that schools will invest in the robot when it becomes commercially available.

Figure 1: The Tesla Optimus robot[5].

In the past robots were mainly used for repetitive, simple tasks in factories. Due to technological advancements, they are now able to navigate themselves effectively in more complex environments. The primary impact of robotics emerged in the 20th century when industrial robots were designed to automate repetitive processes. Robots were intended to take over tasks that are dull, dangerous, and dirty. But with recent advancements and the integration of AI

[1] https://www.tesla.com/we-robot
[2] In reality the Optimus robot relied on tele-ops (human intervention). https://techcrunch.com/2024/10/14/tesla-optimus-bots-were-controlled-by-humans-during-the-we-robot-event/
[3] (Watters, 2023)
[4] https://www.forbes.com/sites/danfitzpatrick/2024/10/12/could-elon-musks-ai-robots-save-a-troubled-education-system/
[5] https://commons.wikimedia.org/wiki/File:Tesla-optimus-bot-gen-2-scaled_(cropped).jpg

systems in robotics, robots are increasingly taking over more and more daily tasks, expanding their capabilities to include a fourth 'D': difficult[6].

As robots enter different environments outside of the industry, such as the work, healthcare, education, and the private home. The application of robots in different domains transforms how we live and how we work. How people react to and act with robots is therefore becoming more important, which is why social sciences and research into human-robot interactions (HRI) are increasingly coming into play in the field of robotics. The interdisciplinary nature of robotics reflects its complexity and broad application, highlighting the importance of striking a balance between the technological innovation of robots and the broad societal and ethical dilemmas involved.

The perceived potential of humanoid robots leads us to this investigatory paper. The presence of a technology like Optimus challenges educational institutions to consider what this technology might mean for their organisations. Should we prepare, not only for the technological capabilities but also for the ethical, social, pedagogical, and logistical challenges of incorporating robots into learning environments?

This report delves into the potential of social robots in tertiary education. The future role of social robots in education is examined by contrasting the ambitious claims of big tech companies with the actual state of the technology. Most importantly this report highlights the importance of responsible and ethical implementation of social robots in education.

[6] https://medium.com/hangartech/robotics-drones-do-dull-dirty-dangerous-now-difficult-a860c9c182a4

Scope of this paper

This paper is aimed to:

- Create a general understanding of educational robots.
- Make readers aware of the possible introduction of social robots in educational contexts given the fast technological developments in robotics.
- Emphasize the importance of responsibly preparing for the application of social robots in tertiary education in the Netherlands.

In addition to literature reviews, the content of this paper is also based on interviews with several experts in the field[7].

[7] Kim Baraka, Cindy Friedman, Eva Leurink, and Inge Molenaar.

1 Terms and definitions

1.1 What is a robot?

Defining a robot proves challenging due to the immense diversity in their functions, structures, and their wide range of purposes. Ask different people what a robot is, and you will likely get different answers. Especially since even roboticists can't agree on a clear definition of a robot[8]. However, there are a few attributes of robots most people agree on: a robot is a machine that can sense what's happening around it, plan what to do, and then take action to get a job done[9]. This makes robots different from ordinary machines or tools, they have a certain degree of autonomy. According to the new European AI-act 'any machine-based system designed to operate with varying levels of autonomy' is an AI system[10]. Since most of today's robots run on AI systems and thus have a sense of autonomy, the line between robots and AI is becoming increasingly blurry. Understanding the role of AI systems in robots is crucial for effectively and responsibly integrating robots into educational contexts.

The merging of AI into robotics marks an important turning point in the field of robotics, enabling robots to exhibit human-like intelligence, human-like movement, and emotional sensitivity. Through AI-driven speech recognition and Natural Language Processing (NLP), robots can engage in natural conversations that go beyond just exchanging information, they can now participate in social small talk or even provide emotional support. Machine learning techniques such as neural networks and deep learning allow robots to learn from data, change their behavior based on experience, and perform real-time decision-making. The integration of human-like intelligence into robots signals a transformative moment in technology: machines are no longer just tools but are becoming companions that engage with us on a more personal, emotional, and interactive level[11].

Nyholm defines a robot as an 'embodied machine with sensors with which it can receive information about the environment, with actuators with which it can respond to their environment, in the service of a certain specified task'[12]. This definition marks the importance of a robot having a body and thus being a physical thing. For the scope of this paper, we will only regard AI systems with a physical body that can sense, plan, and act on their own and, therefore, have a certain level of autonomy, as a robot. There are some nuances considering 'robots' in virtual space. Is an AI represented by a virtual body in an environment where humans are also represented by a virtual embodiment that different from a robot?

1.2 What is a social robot?

A social robot is a robot created to interact with people in interpersonal ways[13]. Often, social robots are designed to be able to interact and communicate with people, following behavioral norms that are typical for human interaction. For example, a social robot may be designed to greet people with a smile, help people with a task, or even provide emotional support. By

[8] (Salvini, et al., 2019)
[9] (Gunkel, 2018)
[10] (European Parliament & Council of the European Union, 2024)
[11] (Prakash et al., 2023)
[12] (Nyholm, 2020)
[13] (Breazeal, 2003)

combining verbal communication with gestures, tone, and sometimes facial expressions, social robots can have intuitive interactions with people. Therefore, social robots are used in domains that require human-robot interactions, such as healthcare or education. According to Fong et al., social robots must be able to communicate with humans, perceive human emotion, possess a distinctive character, understand social aspects of humans, be able to learn and develop social skills and be able to establish and maintain social relationships (2003)[14].

It can be difficult to recognize a social robot, and not everyone agrees on what exactly a social robot is. Furthermore, social robots have the requirement of having a physical embodiment, which spans a broad spectrum of many different physical forms of social robots. There are social robots that are animal-inspired, such as PARO, social robots that are biological-inspired such as the robotic flower, social robots that are integrated into everyday objects such as Alexa or, social robots that mimic animal group behavior called swarm robots, such as the RoboBees (Figure 2). All these social robots have a few things in common: they are all designed to perform social tasks and differ from virtual agents or avatars in their embodied physical appearance[15]. Besides the above examples, a lot of further subcategories and other robots exist that could arguably be put on the spectrum of social robotics. For a more elaborative clarification of the categorization of social robots, refer to the paper of Baraka et al. (2020)[16].

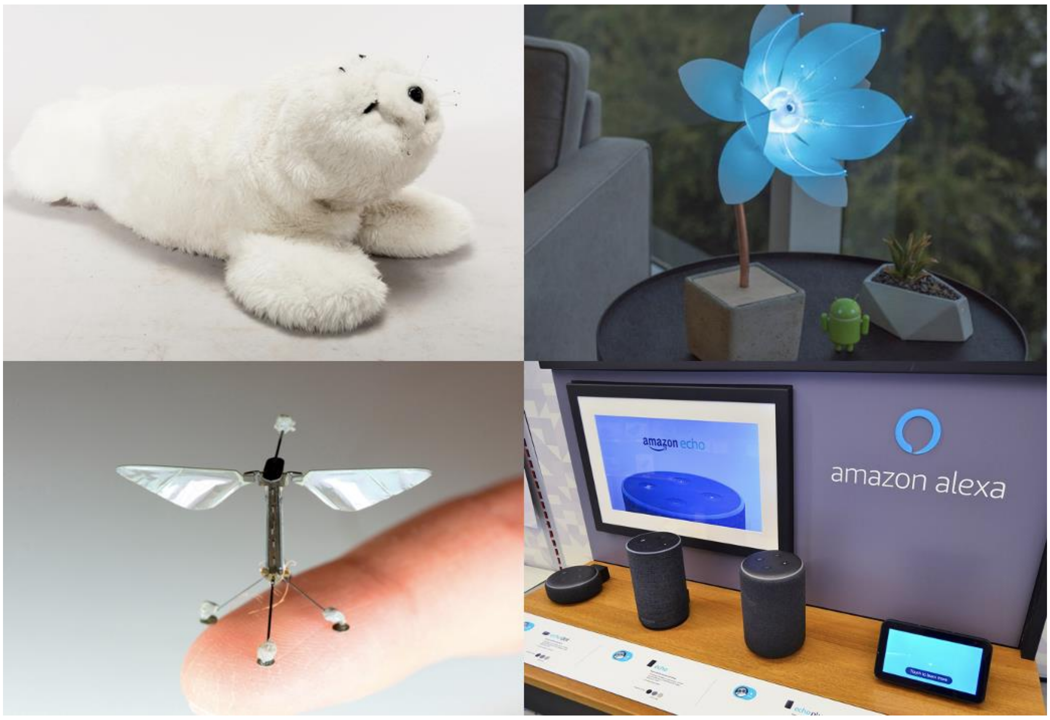


Figure 2: Examples of social robots: Paro, the robotic flower, RobotBees and, Amazon Alexa[17].

[14] (Fong et al., 2003)
[15] (Konijn et al., 2020)
[16] (Baraka et al. 2020)
[17] Paro: https://commons.wikimedia.org/wiki/File:Robots%C3%A4len_Paro_TEKS0057912.jpg
The robotic flower: https://www.roboticgizmos.com/android-things-robotic-flower/
RobotBee: https://www.nationalgeographic.com/culture/article/could-robot-bees-help-save-crops
Amazon Alexa: https://en.wikipedia.org/wiki/Amazon_Alexa#/media/File:AmazonAlexaBooth.jpg

1.3 What is a humanoid robot?

A humanoid robot is a social robot that is designed to look and behave like a human being[18]. Humanoid robots are created to be as realistically humanlike as possible, designed with the intention for them to possibly be mistaken as human beings[19]. Sophia for example, is a humanoid robot designed by Hanson Robotics and can perform many different facial expressions due to her human-like face made of a material mimicking human skin (Figure 3). Recently developed humanoid robots such as Figure02, designed by OpenAI, and NEO beta, designed by 1XTechnologies, are less explicit in their facial expressions and therefore lack certain social interaction features. However, these humanoid robots are designed for their mobility and dexterity, mimicking posture and body movements that are highly human-like. An advantage of humanoid robots like these is their ability to use infrastructures designed for humans[20]. Their appearance for instance resembles a robotic exoskeleton, hiding their mechanical build-up. Figure02 is designed to perform operational capabilities to work in industries requiring physical labor such as warehouses or factories. NEO beta is primarily designed for its application in the household. For both humanoids we can't rule out their possible future applications in other contexts, such as education for example.

Figure 3: Example of humanoid robots: Sophia, Figure01 and NEO beta[21].

In summary, robots have been around for a while. What exactly a robot is, is topic of discussion. Social robots stand out for their potential to engage with humans in meaningful ways. The integration of AI and physical embodiment highlights their transformative potential across various domains. For this paper, we focus on the application of social robots in the education domain for tertiary education, exploring their capabilities, limitations, and future possibilities in enhancing learning experiences.

[18] (Nyholm, 2020)
[19] (Friedman, 2022)
[20] (Nagenborg, 2018)
[21] Sophia: https://www.hansonrobotics.com/sophia/
Figure01: https://www.figure.ai/
NEO beta: https://www.1x.tech/discover/announcement-1x-unveils-neo-beta-a-humanoid-robot-for-the-home

2 Social robots in education

2.1 The potential of social robots in education

Recent research in the field of robotics underscores the potential role of social robots in educational contexts. These robots offer unique opportunities to enhance both teaching and learning experiences with their ability to provide personalized, one-on-one instruction, standing out as key advantage[22]. Another notable advantage of social robots in education is their ability to reduce the workload of human teachers. The meta-analysis of Belpaeme et al. has proved that social robots have the potential to improve learning performance up to the level of human educators (2018)[23]. Limited budgets often restrict the level of individualized attention educators can provide to their students. Social robots offer a cost-efficient alternative, capable of delivering personalized instruction and tutoring for small groups. This not only enhances educational outcomes but also frees teachers to focus on their unique strengths. Besides, robots are patient in rehearsing, and address each user the same[24]. As education continues to evolve to meet the growing demand for personalized and inclusive learning, social robots emerge as a powerful tool for innovation. Their ability to complement and support traditional teaching methods positions them as a valuable asset in classrooms.

Ružić & Balaban have investigated the multifaceted use of social robots in primary and secondary education, which highlights the redundant use of the social robots NAO and Pepper (2024)[25]. NAO and Pepper are both humanoid robots, developed by SoftBank Robotics (Figure 4). Both are widely used in educational research and can adapt to individual learner needs by assessing their user's abilities, attention span, and engagement. NAO and Pepper must be pre-scripted in advance setting out what the robot should say, how it should move, and how it should respond to certain inputs.

Figure 4: Example of educational robots: NAO and Pepper[26].

[22] (Donnnerman et al., 2020)
[23] (Belpaeme et al.,2018)
[24] (Konijn et al., 2020)
[25] (Ružić & Balaban, 2024)
[26] NAO: https://commons.wikimedia.org/wiki/File:NAO_Evolution_.jpg
Pepper: https://commons.wikimedia.org/wiki/File:Pepper_the_Robot.jpg

The broad field of educational robots has already shown their promising impact with the following functionalities:

- **Enhancing learning experiences** in both familiar and novel subjects[27]. Supporting educational institutions by modernizing teacher methods and creating dynamic learning environments.
- **Facilitating language acquisition** by teaching vocabulary, aiding conversations, and practicing pronunciation[28]. Enabling educational institutions to assure linguistic diversity.
- **Improving learners' knowledge and confidence** in subjects like algebra, geometry, and trigonometry[29]. Fostering critical thinking and promoting problem-solving skills.
- Offering essential **support** by facilitating communication, enhancing social skill development, and assisting with emotional regulation, particularly benefiting children with autism spectrum disorder[30]. Making educational institutions more inclusive for students with special needs.
- Offering **emotional support to students** facing stress or loneliness[31]. Supporting educational institutions in student guidance.
- **Practicing soft skills** such as collaborative work and negotiation skills[32]. Contributing to preparing students for professional and social success in the workforce.

While most of the above research in educational robotics has focused on primary or secondary education, or domain-specific education such as STEM courses (Science, Technology, Engineering, and Mathematics) and special education (particularly children with language deficiencies or concentration deficiencies)[33], research on the potential of social robots in tertiary education remains underexplored. Moreover, research is often executed in controlled environments with minimal involvement of human teachers[34]. Fortunately, there is emerging research investigating the use of social robots in tertiary education. The following case study investigates the use of Pepper in university education.

[27] (Baxter et al., 2017)
[28] (Sisman et al., 2018)
[29] (Ahmad et al., 2020)
[30] (Ružić & Balaban, 2024)
[31] (Escobar-planas et al., 2022)
[32] (Alnajjar et al., 2021)
[33] (Sondererger et al., 2022)
[34] (Woo et al., 2021)

Case-Study: How Pepper can provide individualized tutoring for exam preparations for university students (Donnerman et al., 2022).

This study, at the University of Wuerzburg, explored the integration of a humanoid robot, Pepper, by Softbank Robotics, into university education as a supplemental tutoring tool. The robot-supported learning environment was designed to provide individualized tutoring for exam preparation and allowed students to engage with the robot voluntarily. Pepper used speech, gestures, and visual cues to deliver structured learning exercises.

The findings indicated a positive reception of the robot-supported tutoring. Students appreciated Pepper's appearance, perceived personality, and ability to provide adaptive feedback based on their responses. The robot's physical presence and interactive nature were reported to enhance motivation, attention, and concentration compared to traditional learning methods, such as self-study or screen-based exercises. Even in comparison to human tutors, participants highlighted unique benefits of learning with a robot. Moreover, exam performance among students who participated in the robot-supported tutoring was higher than those who did not, demonstrating the effectiveness of the robot. Overall, the study concluded that robot-supported tutoring offers significant potential for enhancing university teaching, increasing engagement, and improving learning outcomes. Students expressed a strong interest in using the learning environment again, underscoring the promise of social robots in higher education.

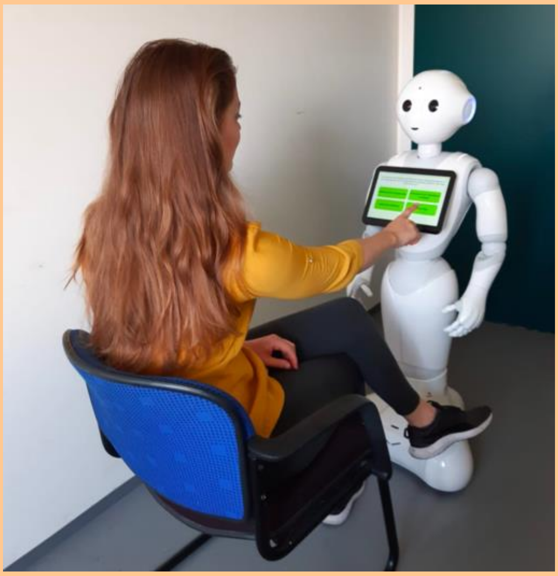


Figure 5: The experimental setup

Social robots hold significant potential to revolutionize education. For example by offering personalized instruction, reducing teacher workload, and supporting diverse learning needs. The application of social robots in tertiary education is beginning to be explored.

2.2 Embodied social robots versus screen-based technologies

Compared with virtual agents or screen-based technologies such as chatbots or personalized learning systems, that are already thoroughly being used in education, embodied social robots offer advantages for their application in education. First, the physical presence of social robots allows for direct interaction with the physical world. Teachers are expected to be able to move through a classroom, to be able to approach a student, and if needed physically manipulate the classroom environment. Although this is not always needed in the context of robotics in education, there are scenarios in which the learning experience benefits from a social robot being embodied and able to autonomously manipulate its direct environment. Second, students show increased learning gains when interacting with a physically embodied social robot, and third, its users show more social behaviors that are beneficial for learning[35]. In the case study described above, it was shown that the physical presence and interactive nature of Pepper enhanced students' motivation, attention, and concentration compared to learning environments with screen-based exercises only[36].

Moreover, the physical presence of the Furhat[37] robot, a social robot with a head only, designed by Furhat Robotics for the purpose of supporting student's mental well-being, proved additional learning gains compared to when a screen-based version of the robot was used[38]. The physical presence of the Furhat robot, elicits cognitive effects in the brain that engage attention, memory, and learning. Furhat is designed minimalistic with a white curved projection surface representing its head (Figure 6). The facial expression of the robot is projected on the projection surface, creating a lifelike experience of the robot's head. The neck of the head is moveable, and its adaptable face allows Furhat to take on different genders, ethnicities, and personalities depending on its application context. Its facial expressions are highly versatile and human-like, creating a sense of relatability and emotional engagement in its interactions with people.

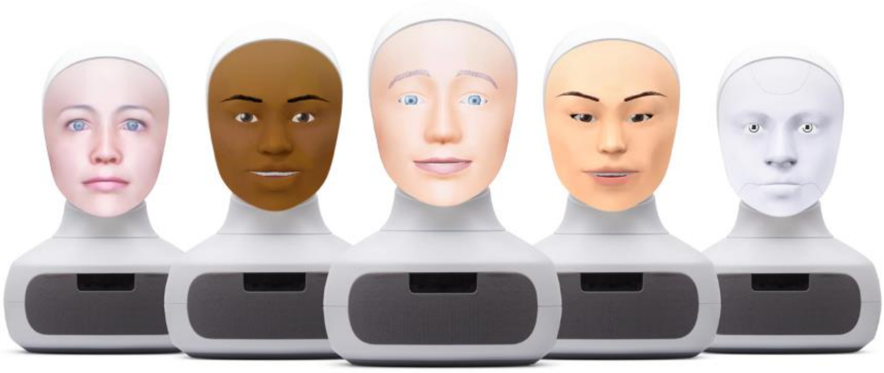


Figure 6: Different embodiments of the Furhat robot[39]

[35] (Berland & Wilensky, 2015)
[36] (Donnermann et al., 2021)
[37] https://www.furhatrobotics.com/use-cases-and-concepts/teacher-robot
[38] (Wedenborn et al., 2019)
[39] https://nl.pinterest.com/pin/music--856739529122887323/

Case-study: How the Furhat robot as a teaching assistant enhances university student learning (Kulathunga, 2023).

This study was conducted as a master thesis to explore the use of Furhat, a social robot, as a teaching assistant in classroom settings at a university. Students interacted with different Furhat personalities to evaluate how these personalities influenced their engagement, attention, and emotional responses during classroom activities. Feedback from the students on these initial interactions revealed that friendly and approachable personalities increased student engagement and elicited positive emotions, while slow speech or unsettling expressions reduced attention and created discomfort. To address these limitations, a new robot personality, Astro Luna, was developed specifically for teaching astrology. Designed to be warm, engaging, and adaptable, Astro Luna aimed to improve interaction and enhance the learning experience. Four participants completed astrology-related tasks with Astro Luna's assistance and provided feedback on their experience.

Figure 7: The different embodiments of AstroLuna.

The findings showed that Astro Luna effectively captured attention and maintained engagement. Students praised its ability to adapt explanations based on their preferences, offering concise or detailed responses as needed. Despite its strengths, some students found the explanations occasionally too lengthy, causing attention to wane. However, the overall response was positive, with the students highlighting Astro Luna's potential to make learning more interactive and engaging. While further refinement is needed, this approach demonstrates the potential of social robots to transform university.

In summary, the physical embodiment of social robots provides distinct advantages over screen-based technologies in education, enhancing student engagement, attention, and learning outcomes. The ability of these robots to directly interact with the environment enriches learning experiences.

2.3 Are we there yet?

The promise we are being sold is that of the Optimus robot of Tesla soon taking a place as a teacher assistant in the classroom. But are we that far yet? After the 'We Robot' event of Tesla, it soon became clear that the very advanced AI-powered humanoid robot Optimus wasn't working as autonomously as Elon Musk made it appear. Most of the robot's actions were actually teleoperated by humans[40]. For example, the conversations that the robots had with people at the event were not held with AI systems or natural language processors, but with real people. This underscores the gap between the marketing promises of big tech companies and the current technological reality[41]. This phenomenon is commonly seen in marketing strategies of big tech companies, where the illusion of automation is used to present technology as advanced, cutting-edge, and attractive[42]. It emphasizes the importance of consumers and investors to critically look towards new technological developments. The technological capabilities of robots often seem very impressive at first, but in practice, they often lack in fulfilling user's expectations. And it is not only the technology itself but also the application of the technology in a certain context - in this case education - that comes with its implications.

Recently developed social robots such as Optimus, NEO Beta, and Figure02 are very sophisticated, and given the fast pace of technological innovation, we can't rule out that these social robots will one day effectively perform in the classroom. However, at this point, there is no evidence that social robots are superior to either human teachers or to other smart technologies in terms of either academic or socio-emotional learning gains[43]. Social robots that are used today in educational settings do not seem to be technologically mature enough to be used in the classroom without additional human support. The humanoid robot Bina48, developed by Hanson Robotics[44], for example, taught a philosophy lecture on ethical reasoning, war theory, and the use of AI in society at West Point Military Academy in 2018 (Figure 8). The humanoid robot surprised many students with its capabilities at first, students were positively engaged and actually took notes[45]. Eventually, the students outperformed the robot cognitively as it wasn't capable of answering certain questions and wasn't able to keep up with the class.

[40] https://www.theverge.com/2024/10/13/24269131/tesla-optimus-robots-human-controlled-cybercab-we-robot-event
[41] https://www.nederlanddigitaal.nl/actueel/nieuws/2024/10/15/nieuwe-tesla-robot-optimus-blijkt-mechanische-turk
[42] Learn more in our (Dutch) SURFshort podcast episode on hype navigation: https://soundcloud.com/surf_short/wat-je-moet-weten-over-hypenavigatie
[43] (Woo et al., 2020)
[44] https://www.hansonrobotics.com/bina48-9/
[45] https://frontcore.com/blog/news/meet-the-first-robot-who-teaches-a-college-course/

Figure 8: BINA48[46]

BINA48 was pre-programmed by AI developers by integrating data relevant to the content of the course. To ensure she relied solely on her preloaded memory rather than external sources, she was disconnected from the internet during the lecture. BINA48 then utilized her stored knowledge and lesson plan to deliver a lecture and responded to student questions based on her programming. Her AI relied on structured datasets and predefined interaction patterns. However, with today's technological advancements, including Large Language Models (LLMs) such as GPT, it is reasonable to expect robots like BINA48 to perform better in classroom settings offering more nuanced discussions, real-time decision-making, and engaging in educational settings.

Despite advancements in social robots, their capabilities still fall short of the promises made by big tech marketing. While robots like BINA48 show potential, current implementations in education rely heavily on human support and pre-programming, highlighting the gap between technological sophistication and practical application in classrooms.

2.4 Robot-teacher collaboration

Social robots can take on different roles in educational settings, changing the setting in which they collaborate with teachers and interact with students. In education, robots can take on the role of teacher (BINA48 example), teacher assistant, tutor, or peer. While the prospect of robots giving a lecture independently and without human involvement might not be very likely, the effective use of social robots as collaborative teacher assistants – or as tutors working alongside teachers in a balanced collaboration – appears more promising[47].

Inge Molenaar's model of 6 levels of automation describes how hybrid human-AI solutions can combine strengths to achieve personalized learning[48]. It is a constructively scaled model in which

[46] https://images.app.goo.gl/YAn7UQE5CprPwyts7
[47] (Belpaeme et al., 2018)
[48] (Molenaar, 2022)

at level one the teacher has full control over the learning experience of students, and level six is the complete opposite where the technology - in this case the AI system for personalized learning - has full control over all teaching tasks (Figure 9). This model can be transposed to hybrid human-robot collaboration to improve teaching quality in the classroom. Moreover, transposing this model of automation to this new context can facilitate discussions about the future role of social robots in tertiary education.

Therefore, suggestions on what teacher-robot collaboration in the classroom could potentially look like for each level of automation are suggested. This will be done by providing examples from previous research and placing these on the levels of automation based on the scenario in which the social robot is used. However, as we progress to higher levels of automation, there are currently no established examples. We address this gap by presenting scenario sketches, illustrating the future potential implementation of social robots in tertiary education.

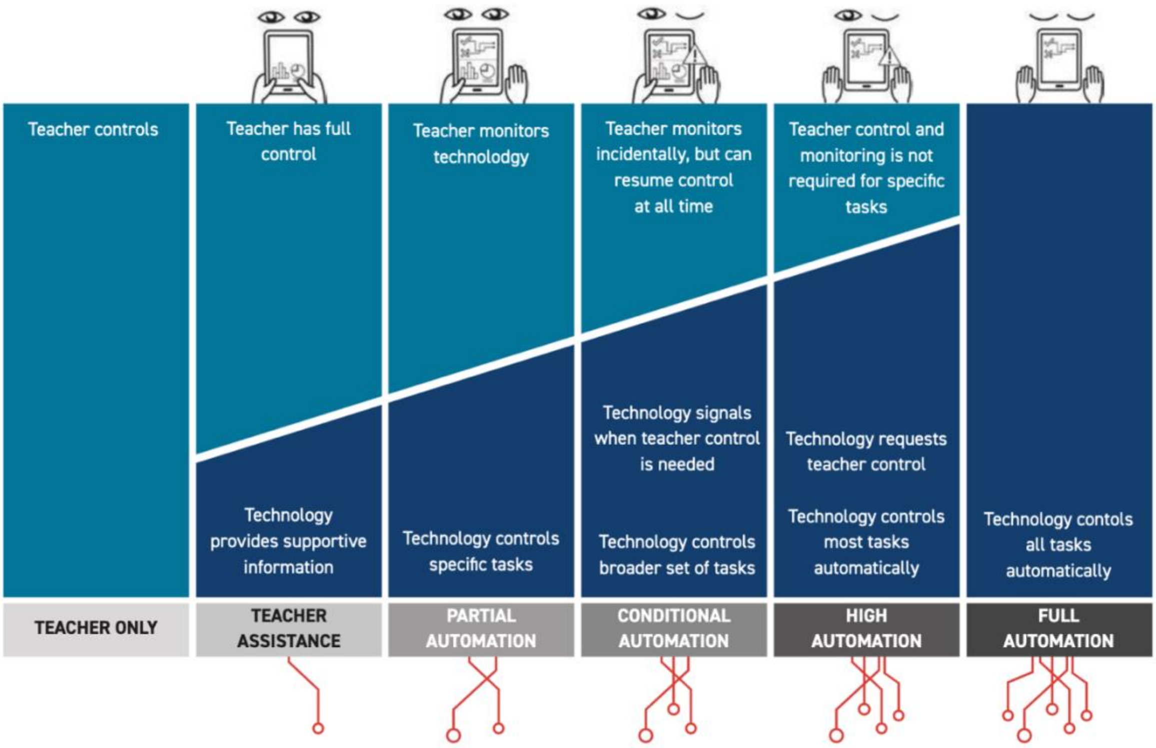


Figure 9: Levels of automation[49].

The first level of automation in the context of human-robot collaboration in education speaks for itself. The teacher is completely in control of the students' learning process and the use of a robot is not utilized in the learning process. In the second level of automation, teacher assistance, the robot supportive information to the teacher or students. The technology provides access to learning materials and offers additional information about learners' activities. The technology does not control any aspects of the learning environment. Social robots in this level of automation could be fulfilling tasks such as repeating teacher instructions to students or providing acces to learning materials. The robot is unable to interact with students or respond to

[49] (Molenaar, 2022)

questions. Here the added value of an embodied social robot compared to a screen-based supporting technology is not sufficient.

In the third level, partial automation, a social robot takes over specific well-defined tasks that are predefined by the teacher. In this level of automation, the teacher is completely in control of the student's learning process, but the teacher outsources a compact, simple task to the robot. An example is the research by Konijn & Hoorn in which pupils (aged 8-10 years) were picked up one-by-one from the classroom to take part in a one-to-one tutoring session with the NAO robot (2020)[50]. The NAO robot was programmed in advance to tutor specific times and deliver repetitive, structured tasks to the pupils. These tasks were tightly defined, with the robot having no autonomy in modifying the learning content during the session.

According to the model, in level three of automation, the robot must be capable of providing immediate feedback to its users. In this research, when pupils were doing good, NAO clapped his hands or cheered "Fantastic" or "Well done." When pupils answered the mathematic question incorrectly, NAO would say: "Too bad, that is wrong, let's try again". Furthermore, the human teacher was still in control of the child's overall learning process, monitoring the child's learning progress, performing all other educational tasks, and making pedagogical decisions. During the process of this research, the interaction between the robot and the child was monitored by the researcher. The research results support the idea of a humanoid robot tutor like NAO as an effective support tool.

In level four, conditional automation, the division between the teacher's and robot's tasks is almost equally divided. The robot performs a cluster of related tasks and is capable of providing step-by-step feedback on the learning content. The robot works autonomously within the defined tasks, however, the robot must be able to recognize and alert the human teacher when its limit is reached and when human involvement is necessary. The robot is not capable of effectively working outside its domain and cannot provide pedagogical or emotional support to students. It's therefore the teacher's role to take over when needed and provide deeper insights where the robot's capabilities are insufficient.

[50] (Konijn & Hoorn, 2020)

Scenario sketch 1: A social robots teaching presentation skills[51]

In a tertiary education scenario focused on presentation skills, a social robot with conditional automation can support students by providing feedback on non-verbal aspects of their presentation, such as posture, gestures, and body language. Equipped with advanced AI, like the Honost Mirror at the University of Utrecht, the robot observes students' movements and offers step-by-step guidance on improving eye contact, stance, and gestures, helping students refine their delivery. It can suggest adjustments in real-time, such as correcting posture or suggesting more expressive hand gestures to emphasize key points, offering a more interactive and personalized learning experience. This suggestive robot's embodiment adds significant value over the (already existing) screen-based AI by enabling real-time demonstrations and hands-on corrections, such as physically guiding the student to adjust their stance or by demonstrating which gestures or body movements could be made at certain moments of the presentation. Its physical presence allows students to receive immediate visual and physical feedback, making the learning experience more immersive.

However, the robot's role is limited to providing feedback on the non-verbal aspects of the presentation. It cannot assess or provide insight into the content of the presentation itself because this ability is not included in the Honest Mirror AI programming. In these cases, the human teacher steps in to review the content, providing valuable feedback on the organization, clarity, and depth of the presentation. The robot autonomously alerts the teacher when a student's verbal content requires attention, ensuring that the teacher remains the primary source of feedback on intellectual and content-based aspects of the presentation. This collaborative approach ensures a balanced and effective learning experience.

In level five, high automation, the primary responsibility for managing and executing the learning process shifts to the robot, with the teacher assuming a supportive role through a limited set of clustered tasks. The robot is therefore capable of providing immediate feedback on specific content and can set personalized goals for students. It's hereby noteworthy that this social robot is capable of teaching without human oversight and will only alert and ask for human intervention in unforeseen circumstances or when problems occur that the robot itself can't solve. The teacher's primary focus becomes integrating the robot into the curriculum and ensuring alignment with educational goals, with minimal monitoring required. Although there is currently no existing example of robots that can operate at this level of automation, a scenario sketch of how this would look can be provided.

[51] Based on Honest Mirror https://husite.nl/digitalehu/dlo-innovatiethemas/showcases-dlo-kwaliteitsafspraken/honest-mirror/

Scenario sketch 2: A social robot teaching a neuroscience course

Imagine a tertiary education scenario where a social robot is integrated into a neuroscience course by serving as both a teacher and facilitator, with minimal need for human oversight. This humanoid robot is designed to deliver personalized and interactive learning experiences, and its humanoid form enables it to interact physically with its students. The robot has a touchscreen interface on its chest allowing students to access lesson materials, quizzes, or topics of interest. The robot combines artificial intelligence, real-time feedback capabilities, and immersive visualization technologies to support complex subjects such as the autonomy of the brain. The robots uses 3D holographic models projecting the human brain, enabling it to guide students through complex topics such as neuronal pathways, brain regions, and their specific functions. Its audio-visual system, ensures clear and engaging communication, while its wheels allow for movement within classrooms or lecture halls. Moreover, its arms allow for pointing to the brainstructures displayed on the holographic model.

The robot stimulates the students to actively learn by asking them to label certain brain regions on the holographic model or by asking students what the function of certain brain regions are. It can help students remember the names and places of brain regions through lots of repetition and interactive quizzes, and is capable of adapting its learning strategies based on the student's need and learning progress. For example, if a student struggles to understand the role of neurons, the robot can animate the holographic model to show a detailed synaptic process. In this highly automated scenario, the teacher’s role shifts toward curriculum integration and minimal monitoring. Teachers focus on ensuring that the robot aligns with the lessons plan and curricular requirements while intervening only in unforeseen circumstances or when technical issues arise. This collaboration between robot and human educators offers a glimpse into the possible future of education.

In level six, full automation, the social robot autonomously manages cross-task domains and situations, entirely replacing the role of human teachers. This level represents a paradigm where social robots monitor the teaching experience without human intervention. While promising for personalized and scalable education, its integration into formal schooling raises questions about feasibility, ethical considerations, and the evolving responsibilities of teachers. Social robots working in level six of automation remain hypothetical due to current limitations in artificial intelligence and robotics. Transitioning to level six would require not only technological advancements but also a significant rethinking of educational structures and teacher roles.

Teacher-robot collaboration holds significant potential in enhancing the educational experience. Given the current state of the technology, human-teacher collaboration is working primarily at level two or three of automation, where robots provide supportive tasks and feedback. Higher levels of automation have not yet been realized, although we cannot rule out the possibility that robot-teacher collaboration may eventually reach higher levels of automation as technology advances.

3 The responsible use of social robots

3.1 Ethical implications

The application of social robots in tertiary education raises a variety of ethical implications, including privacy concerns, data security, bias in AI, accessibility, and the potential for job displacement[52]. However, for the purpose of this paper, we will focus specifically on the ethical implications related to the embodiment aspects of social robots, such as their physical presence, autonomy, and social interactions with students. These so called 'soft risks' are unique to robots with embodied forms and have distinct ethical considerations compared to other technologies. While privacy and data security are common concerns across many technological domains, the way social robots engage with students through human-like behaviors and physical presence introduces additional challenges related to emotional manipulation, dependency, and social dynamics that warrant closer examination.

As robots' autonomy and the complexity of their interactions with students in an educational context increases, ethical questions arise regarding who is in control and who is accountable when something goes wrong. The causal chain leading to damage or conflict is not clearly recognizable[53]. It is often unclear who can be considered responsible for the unforeseen implications of using robots. While the teacher or educational institutions often seem to have authority over the robot's operations, its behavior is ultimately shaped by decisions made during its design and programming. This is also called the problem of many hands: the difficulty of holding an individual accountable because the responsibility is distributed over many different individuals[54]. "With distributed agency comes distributed responsibility"[55]. Therefore, the responsibility for the usage of social robots in education lies both in its developers, its users, the manufacturers, and educational institutions incorporating social robots in their curriculum, but also in policymakers and the student's parents.

The design of social robots makes them appear as more human-like, mimicking behaviour simulating emotions, empathy, moral reasoning, an reciprocity [56]. However, these interactions are inherently deceptive[57] as robots cannot truly engage in reciprocal relationships [58]. Highlighting the fundamental limitations of human-robot relationships despite their human-like appearance. Establishing social bonds with robots can have positive effects but can also entail major risks. People might form emotional bonds with robots and whether this improves quality of life or just creates dependencies is not clear-cut and depends on the context[59]. In the context of tertiary education, students might start to overly rely on robots for their learning processes and become attached to or even dependent on them. This can impair their ability to form skills essential for academic self-development or impair their independent working abilities.

[52] (Müller, 2023)
[53] (Matthias, 2004)
[54] (Nissenbaum, 1996)
[55] (Taddeo & Floridi, 2018)
[56] (Friedman, 2022)
[57] (Nyholm et al., 2023)
[58] (van Wynsberghe, 2022)
[59] (de Graaf, 2016)

The process of teaching and learning relies heavily on a teacher's ability to deliver content. In addition, it's also necessary for a teacher to ensure that learning is taking place by actively evaluating students' understanding. Teachers must possess the cognitive skill to 'read the room', checking whether students are engaging with, and understanding the learning material. However, at this point, robots lack human judgment, common sense, appreciation of the larger picture, understanding of the intentions behind people's actions, and understanding of values and anticipation of the direction in which events are unfolding[60]. These human attributions ensure that no matter how sophisticated robots become in the future, robots will not be able to completely replace teachers without major changes in teaching methods and the current structure of the educational system[61]. Yes, robots are becoming more sophisticated in performing specific human-like behavior, but they cannot fully understand what it is like to be a human, feel human emotions, and make moral decisions. The robot is deceptive, but the feeling it elicits in its user is real. This, however, raises an important question about whether such human-like qualities are necessary for effective teaching in tertiary education.

3.2 What can we already do to prepare?

The possible integration of social robots into Dutch follow-up education could be of great potential: transforming classrooms into a space of personalized learning, where embodied robots deliver socially interactive teaching, offering undivided attention to each student, and having infinite patience. However, as we consider introducing these technologies into the classroom, we must ask the question: how can we ensure social robots are used responsibly, ethically, and in alignment with public values? Social robots will shape future educational environments in ways we cannot foresee at this point. We must be prepared to avoid a purely technological-drive use of robots without critical reflection on pedagogical and ethical implications, as well as implementation risks. But where do we start?

As the famous Collingridge-dilemma clearly describes, it is hard to effectively intervene in the development of new technologies[62]. At early stages of development, there is a lack of understanding of the technology's future impact, at later stages the technology is too embedded in our society to be easily altered. The Collingridge-dilemma has implications for the social control of innovation of social robots. As discussed earlier, the field of social robotics is changing rapidly, and its technological refinement is enabling robots to effectively change educational settings. This underscores that technologies are not neutral, they embed values that affect human behavior and societal norms[63]. A social robot designed to encourage students in a classroom does more than simply assist the teacher. It actively transforms the way students learn, reshapes the classroom experience, and influences the teacher's approach to teaching. But how can we ensure the responsible use of social robots in education?

According to Responsible Research Innovation (RRI) theories, responsible technologies are achieved through collaboration between stakeholders that innovate ethically, empowering them to work together to shape a responsible technological future[64]. In the context of educational

[60] (Sharkey, 2017)
[61] (Torras, 2023)
[62] (Collingridge, 1982)
[63] (Verbeek, 2016)
[64] (Salvini et al., 2019)

social robots, the scope of relevant stakeholders for innovating extends beyond those responsible for technological innovations and those directly impacted by the technology. It also includes educational institutions, student's parents, policymakers, robot developers, manufacturers, robotic engineers, and the end users of the technology. The second part of Responsible Research Innovation is innovating ethically. SURF's responsible technology model can provide direction here, offering actionable recommendations (Figure 10)[65]. When innovating ethically, it is useful to see responsibility as a practice; there is not a fixed outcome on how to deal with social robots in education, we should continuously reflect and adapt on the capabilities and implications of social robots as their technology continuously evolves. This we must do, in an ongoing process of 'doing' ethics, by embracing ethics as an evolving process of engagement with technology's real-world effects.

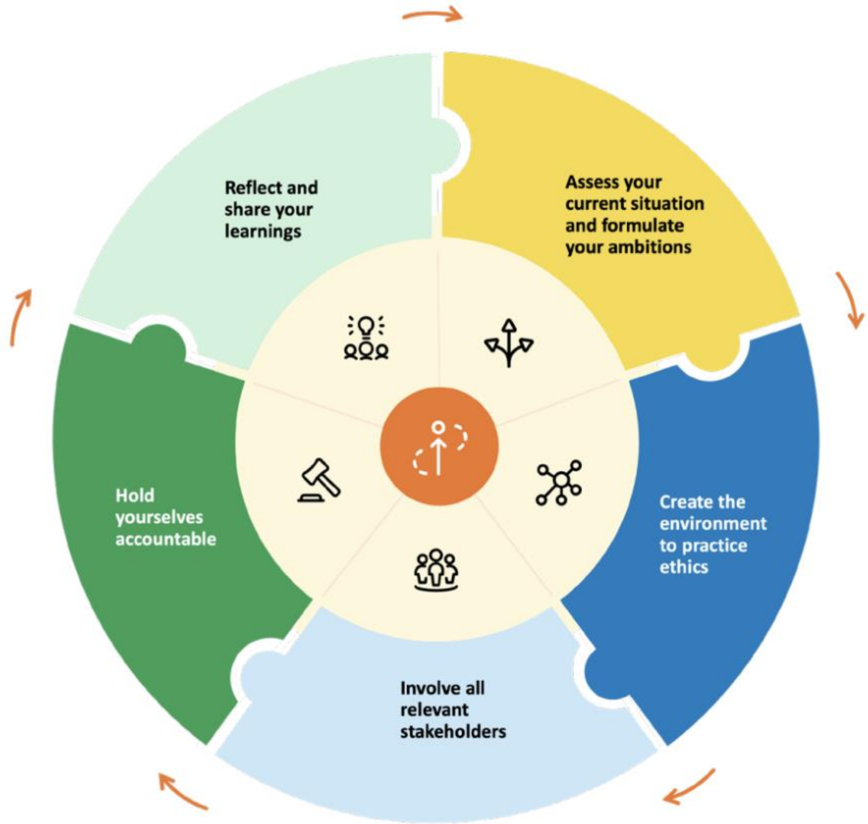


Figure 10: The Responsible Technology model[66]

Moreover, when 'doing' ethics, it is important to be aware that our traditional ethical standards relating to human-human interactions will not naturally transfer to the ethics of human-robot interaction[67]. We need to continuously update our current ethical frameworks to address the new ethical issues that arise with new forms of human-robot interaction in different educational contexts. Drawing from expert interviews, it can be stated that as robots shift from being seen as tools to companions, greater focus is required on ethical frameworks that address the 'soft risks'

[65] (SURF, 2023)
[66] (SURF, 2023)
[67] (Nyholm, 2020)

of robotics, including emotional manipulation, dependency, attachment, and other social dynamics associated with human-robot interactions.

Besides, not only the technology itself but also its implementation in different contexts brings implications. Integrating social robots into educational practices depends on factors such as technological infrastructure, budget constraints, and differing educational philosophies[68]. Where schools with advanced technology and funding can adopt robots more easily, under-resourced institutions risk being left behind. Additionally, achieving effective teacher-robot collaboration at high levels of automation requires restructuring educational systems and redefining teacher roles. This adoption also rests on teaching methodologies: student-centered schools may embrace robots, whereas schools adopting a more traditional way of teaching may resist. Furthermore, reliance on social robots could create vendor lock-in, limiting schools to a single provider's technology and reducing their ability to adapt to future needs.

[68] (Ružić & Balaban, 2024)

Conclusion

Social robots have the potential to transform tertiary education, offering personalized learning experiences, supporting diverse educational needs, and reducing teacher workloads. Their physical embodiment provides an advantage over screen-based technologies by enhancing student engagement, attention, and learning outcomes. However, the implementation of social robots in education comes with challenges and implications that must be addressed responsibly. As social robots shift from being perceived as tools to companions, ethical frameworks must adapt to address the 'soft risks' associated with human-robot interaction, such as emotional manipulation, dependency, and attachment.

Moreover, as technological advancements continue to shape society, the integration of social robots into education is plausible. While big tech companies make big promises, reality shows us that robots still need substantial refinement to effectively support learning in tertiary education. Additionally, their successful adoption depends not only on technological progress but also on transforming the context in which they are used. Traditional teaching methods and teacher roles must evolve alongside these technologies.

Resistance to social robots is natural, as people often hesitate to adopt new technologies. Still, educational institutions must be prepared to navigate this transition responsibly when social robots enter tertiary education. Stakeholders need to work together in shaping a responsible technological future. Staying ahead of a trend means staying on top of the state of the art and remaining flexible to avoid vendor lock-in. Institutions can guarantee this by already thinking about possible redefinitions of teacher roles, and the possible restructuring of curricula. This involves recognizing the importance of ongoing ethical reflection and adaptation while ensuring that robots enhance the educational experience and align with public values.

Bibliography

Ahmad, M. I., Khordi-moodi, M., & Lohan, K. S. (2020). Social Robot for STEM Education. ACM/IEEE International Conference on Human-Robot Interaction, (pp. 90–92). https://doi.org/10.1145/3371382.3378291

Alnajjar, F., Bartneck, C., Baxter, P., Belpaeme, T., Cappuccio, M. L., Di Dio, C., Eyssel, F., Handke, J., Mubin, O., Obaid, M., & Reich-Stiebert, N. (2021). Robots in education. Routledge. https://doi.org/10.4324/9781003142706

Baraka, K., Alves-Oliveira, P., & Ribeiro, T. (2020). An extended framework for characterizing social robots. In Human-robot interaction: Evaluation methods and their standardization, (pp.21–64). Springer. https://doi.org/10.1007/978-3-030-42307-0_2

Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017). Robot education peers in a situated primary school study: Personalisation promotes child learning. PLOS ONE, 12(5), e0178126. https://doi.org/10.1371/journal.pone.0178126

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Science Robotics, 3(21). https://doi.org/10.1126/scirobotics.aat5954

Berland, M., & Wilensky, U. (2015). Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. Journal of Science Education and Technology, 24(5), (pp. 628–647). https://doi.org/10.1007/s10956-015-9552-x

Breazeal, C. (2003). Toward sociable robots. Robotics and Autonomous Systems. https://doi.org/10.1016/s0921-8890(02)00373-1

Collingridge, D. (1982). The social control of technology. Frances Pinter.

Donnermann, P., et al. (2020). Integrating a social robot in higher education – A field study. IEEE Conference Publication. https://doi.org/10.1109/RO-MAN47096.2020.9223602

Escobar-Planas, M., Charisi, V., & Gomez, E. (2022). "That robot played with us!" Children's perceptions of a robot after a child-robot group interaction. Proceedings of the ACM on Human-Computer Interaction, 6(CSCW2), (pp. 1–23). https://doi.org/10.1145/3555118

European Parliament & Council of the European Union. (2024, June 13). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending various regulations and directives (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. Robotics and Autonomous Systems, 42(3), (pp.43–166). https://doi.org/10.1016/S0921-8890(02)00373-1

Friedman, C. (2022). Ethical concerns with replacing human relations with humanoid robots: An ubuntu perspective. AI and Ethics, 3(2), (pp.527-538). https://doi.org/10.1007/s43681-022-00186-0

de Graaf, M. M. (2016). An ethical evaluation of human–robot relationships. International Journal of Social Robotics, 8(4), (pp. 589–598). https://doi.org/10.1007/s12369-016-0368-5

Gunkel, D. (2018). Robot rights. MIT Press.

Konijn, E. A., Smakman, M., & van den Berghe, R. (2020). Use of robots in education. In The International Encyclopedia of Media Psychology, (pp. 1-8). https://doi.org/10.1002/9781119011071.iemp0318

Konijn, E. A., & Hoorn, J. F. (2020). Robot tutor and pupils' educational ability: Teaching the times tables. Computers & Education, 157(103970). https://doi.org/10.1016/j.compedu.2020.103970

Kulathunga, D. (2023). Exploring the experience of different robot personalities in enhancing university students' learning (Master's thesis, Tampere University). Tampere University Repository. https://trepo.tuni.fi/bitstream/handle/10024/157790/KulathungaDakshika.pdf?sequence=2

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and Information Technology, 6(3), (pp. 175–183). https://doi.org/10.1007/s10676-004-3422-1

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. European Journal of Education, 57(4), (pp. 632–645). https://doi.org/10.1787/589b283f-en

Müller, V. C. (2023). Ethics of artificial intelligence and robotics. In The Stanford Encyclopedia of Philosophy (Fall 2023 Edition). https://plato.stanford.edu/archives/fall2023/entries/ethics-ai/

Nagenborg, M. (2018). Urban robotics and responsible urban innovation. Ethics And Information Technology, 22(4), (pp. 345–355). https://doi.org/10.1007/s10676-018-9446-8

Nissenbaum, H. (1996). Accountability in a computerized society. Science And Engineering Ethics, 2(1), 25–42. https://doi.org/10.1007/bf02639315

Nyholm, S. (2020). Humans and robots: Ethics, agency, and anthropomorphism. Rowman & Littlefield.

Nyholm, S., Friedman, C., Dale, M. T., Puzio, A., Babushkina, D., Löhr, G., Gwagwa, A., Kamphorst, B. A., Perugia, G., & IJsselsteijn, W. (2023). Social Robots and Society. In Open Book Publishers (pp. 53–82). https://doi.org/10.11647/obp.0366.03

Prakash, N., Atiq, A., Shahid, M., Rani, J., & Dikshit, S. (2023). Merging minds and machines: The role of advancing AI in robotics. EAI Endorsed Transactions on Internet of Things, 10. https://doi.org/10.4108/eetiot.4658

Ružić, I., & Balaban, I. (2024). The use of social robots as teaching assistants in schools: implicatons for research and practice. Revista de Educación a Distancia, 24(78). https://doi.org/10.6018/red.600771

Salvini, P., Palmerini, E., & Koops, B. (2019). Robotics and responsible research and innovation. In Edward Elgar Publishing eBooks, (pp. 407–424). https://doi.org/10.4337/9781784718862.00037

Sharkey, A. (2017). Can we program or train robots to be good? Ethics and Information Technology, 22(4), (pp. 283–295). https://doi.org/10.1007/s10676-017-9425-5

Sisman, B., Gunay, D., & Kucuk, S. (2018). Development and validation of an educational robot attitude scale (ERAS) for secondary school students. Interactive Learning Environments, 27(3), (pp. 377-388). https://doi.org/10.1080/10494820.2018.1474234

Sonderegger, S., Guggemos, J., & Seufert, S. (2022). How social robots can facilitate teaching quality: Findings from an explorative interview study. In Lecture Notes in Networks and Systems (pp. 99–112). Springer. https://doi.org/10.1007/978-3-031-12848-6_10

SURF. (2023). Responsible tech: On public values and emerging technologies. Utrecht. (Authors: Baten, D., Walker, J.) https://doi.org/10.5281/zenodo.10054653

Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. Science, 361(6404), (pp. 751–752). https://doi.org/10.1126/science.aat5991

Torras, C. (2023). Ethics of social robotics: Individual and societal concerns and opportunities. Annual Review of Control Robotics and Autonomous Systems, 7(1), (pp.1–18). https://doi.org/10.1146/annurev-control-062023-082238

Verbeek, P.P. (2016). Toward a Theory of Technological Mediation. Technoscience And Postphenomenology: The Manhattan Papers (pp. 189–204). Lexington Books. https://ris.utwente.nl/ws/files/21754033/theory_of_mediation.pdf

Watters, A. (2023). Teaching machines: The history of personalized learning. MIT Press.

Wedenborn, A., Wik, P., Engwall, O., & Beskow, J. (2019). The effect of a physical robot on vocabulary learning. arXiv.org. https://arxiv.org/abs/1901.10461

Woo, H., LeTendre, G. K., Pham-Shouse, T., & Xiong, Y. (2021). The use of social robots in classrooms: A review of field-based studies. Educational Research Review, 33, 100388. https://doi.org/10.1016/j.edurev.2021.100388

Van Wynsberghe, A. (2022). Social robots and the risks to reciprocity. AI & Society, 37(2), (pp. 479–485). https://doi.org/10.1007/s00146-021-01207-y

Colofon

Contact us at responsible-tech@surf.nl

Authors
Merel van Steenis, intern at SURF
Under supervision of Duuk Baten and John Walker

Thankful of the input of
Kim Baraka
Cindy Friedman
Eva Leurink
Inge Molenaar

Visualisations
Cover page: Created using Midjourney using prompt 'A robot in between two humans speaking about education in library. corporate image. Pen drawing --ar 16:9 --v 6.1'

Figure 1: https://commons.wikimedia.org/wiki/File:Tesla-optimus-bot-gen-2-scaled_(cropped).jpg
Figure 2: Paro: https://commons.wikimedia.org/wiki/File:Robots%C3%A4len_Paro_TEKS0057912.jpg
The robotic flower: https://www.roboticgizmos.com/android-things-robotic-flower/
RobotBee: https://www.nationalgeographic.com/culture/article/could-robot-bees-help-save-crops
Amazon Alexa: https://en.wikipedia.org/wiki/Amazon_Alexa#/media/File:AmazonAlexaBooth.jpg
Figure 3: Sophia: https://www.hansonrobotics.com/sophia/
Figure01: https://www.figure.ai/
NEO beta: https://www.1x.tech/discover/announcement-1x-unveils-neo-beta-a-humanoid-robot-for-the-home
Figure 4: NAO: https://commons.wikimedia.org/wiki/File:NAO_Evolution_.jpg
Pepper: https://commons.wikimedia.org/wiki/File:Pepper_the_Robot.jpg
Figure 5: (Donnerman et al., 2022)
Figure 6: https://nl.pinterest.com/pin/music--856739529122887323/
Figure 7: (Kulathunga, 2023).
Figure 8: https://images.app.goo.gl/YAn7UQE5CprPwyts7
Figure 9: (Molenaar, 2022)
Figure 10: (SURF, 2023)

Preferred citation
SURF (2025). Social Robots in education. Should we prepare for social robots in education?. Utrecht. (Author: van Steenis, M.)

This exploratory paper is written as part of an internship assignment at SURF

03-03-2025