Analysis of human-robot interaction metrics

by

Nathan Kisjes

A thesis presented for the bachelor's degree of
Life Science and Technology

Under the supervision of Deniz Başkent and Gloria Araiza Illan
Seccond assessor Bart Verkerke

rijksuniversiteit
groningen

Faculty of Science and Engineering

June 2021

1 summary

In the last decades, lots of new technology in robotics has been developed. As a result more and more people interact with robots in their everyday life. Therefore it has grown important to study how humans interact with robots. This paper presents a brief history of robotics, and introduces the field of Human-Robot Interaction (HRI). Which concerns all physical, psychological, social and cultural factors in regard to their interaction with human users. It is a multidisciplinary field bringing together engineers, designers, philosophers and sociologists to study and improve how humans interact with robots.

Research on HRI can be subdivided into three categories. User oriented research, robot oriented research and system oriented research. These three approaches, and in what kind of research they are implemented are explained. Along with their advantages and disadvantages and examples in recent studies.

Furthermore, different methods of data acquisition commonly used in HRI research are explained. These are task performance metrics, self assessment metrics, observational metrics and psycho-physiological metrics.

Finally, 12 recent studies are compared, highlighting examples from different categories of HRI research and different metrics used in data collection

Table of Contents

Page

1 summary **1**

2 Introduction **3**
2.1 History of robotics . . . 3
2.2 Different tasks robots are used for today . . . 3

3 Human-Robot Interaction (HRI) **4**
3.1 Evaluating HRI . . . 4
3.1.1 Approach to analyse HRI . . . 4
3.1.1.1 User oriented research . . . 5
3.1.1.2 Robot oriented research . . . 5
3.1.1.3 System oriented research . . . 5
3.1.2 HRI data collection . . . 6
3.1.2.1 Task performance metrics . . . 6
3.1.2.2 Self assessment metrics . . . 6
3.1.2.3 Observational metrics . . . 6
3.1.2.3.1 Video analysis . . . 7
3.1.2.4 Psycho-physiological metrics . . . 7
3.2 Comparison of recent studies . . . 7

4 Conclusion **12**

References **13**

2 Introduction

2.1 History of robotics

It is no recent development that humans fantasize about human-like machines to assist us in our day to day life. Across many cultures, stories like this can be found that describe automated lifelike tools. Ancient Greek myths talked about golden mechanical maidens forged by the gods, who appeared like humans and could speak and help the gods with all their tasks [7].

It was in 1921 when the Czech Karel Capek first coined the word robot, derived from the Czech word robota meaning forced worker, in the play Rossum's Universal Robots [10]. It was a story fantasising about automated machines, with capabilities that were incredible at the time, eventually rising up against their human makers.

It still took a long time before these automated machines became a reality. The first industrial robot was developed in the 1950's: a mechanical arm called Unimate, working on an automotive assembly line [6], where it had the job of welding die castings to autobodies. This process emitted harmful gasses that were dangerous for the workers previously performing the job.

Since the conception of robots, people have had the desire to create robots that look and function like humans do, recreating physical human attributes like walking on two feet or robotic hands to grasp objects. Since the 1990's, the field of social robotics took off, robots that communicate in ways natural to us humans, following social rules relevant for its task. Treating humans not just as physical objects with a given size and location, but as communicative partners with unpredictable behaviours to respond to [25]. Some early examples of social robots with communicative capabilities and or voice input are MAIA [3], RHINO [2], and AESOP [14]. These are three systems used for very different tasks that all have some basic form of communication. MAIA was developed to navigate through a room and hand objects to the user, activated by voice. RHINO was a robot designed to give guided tours through a museum, while safely navigating through crowds of people. AESOP was a voice activated robotic arm to help surgeons during laparoscopic surgeries. At the time, these robots were state of the art. However, their qualities were far from replacing the human alternative. For example, using AESOP increased the time it took to perform the surgery, and greatly decreased the comfort of the surgeon. This was partly due to the physical capabilities of the robot, but also because of the limited verbal command inputs by the surgeon.

2.2 Different tasks robots are used for today

Nowadays, robots are no longer science fiction but part of everyday life. They have become widespread in many facets of our lives. Their use in surgery, industry, household tasks, education and health care means many people interact with robots throughout their day.

To highlight an example, recently a growing interest has developed to assist medical staff in hospitals and other care facilities. Robotic medical assistants such as Moxi [16] and TUG [20] can help perform non-critical tasks, such as retrieving supplies. Which allows nurses and other healthcare workers more time for direct patient care. The demand for such healthcare robots is highlighted in the recent COVID-19 pandemic, which showed the effect of shortages of medical staff.

Another example of a growing field of robotics is that of commercial household robots. Us humans have always searched for ways to relieve ourselves from our daily chores. And one way to do this is by the use of robots to do these tasks for us. More and more people use robotic vacuum cleaners that autonomously navigate the room and clean every bit of the floor.

With robotics more intertwined with our day to day life then ever, we also interact with these robots more than we have before. The question arises, how can interactions between humans and robots be analysed and what approaches can be taken to study it?

3 Human-Robot Interaction (HRI)

The field of Human-Robot Interaction concerns physical, psychological, social and cultural factors of the design, implementation and evaluation of robotics in regard to their interaction with human users. Creating a successful HRI requires collaboration from a variety of domains. Engineers, designers, philosophers and sociologists need to work together to develop hardware and software and analyse its effect on the humans that interact with it. It regards the robot's physical embodiment, behaviour, control interface, feedback and manipulation of the outside world [4].

An area of robotics that has always been of interest is the use of robots as a conversational partner. This is an fascinating topic in regard to HRI as the easiest way of communication between humans is through speech. No instructions are needed when the interface of the robot is through verbal commands. This means no training is needed to interact with the robot, and nearly everyone is able to operate it, making it attractive for many uses: e.g. education, healthcare, public service. As verbal communication is so incredibly complex, lots of research is done to study robots with speech capabilities. Trying to create a more natural way of communicating with social robots. An example of this is [9], studying the effect of empathetic voice intonation in healthcare robots. It's results show a preference for healthcare robots that speak using a more empathetic, human-like voice.

While it is currently a very popular sub topic, the field of HRI is not limited to studying verbal communication with humanoid robots. It concerns all interactions one could possibly have with robots. This includes all physical and sensory interactions between the two. An example of another domain in HRI is the physical control interface of a robot. For example, between a surgeon and a surgical robot. For the surgeon-robot system to perform the surgery well, the surgeon has to be able to control the robot with extreme precision and accuracy to make sure no surrounding tissues are harmed. The feedback from the robot has to be quickly and clearly brought back to the surgeon.

With the use of some surgical robotics, the surgeon is not operating on the patient directly. This means they are not able to see and feel the result of their actions. It is therefore necessary that the robot also provides sensory feedback to the surgeon. New systems are being developed to provide the surgeon with this information. One example of this is (put reference Force sensing and control for a surgical robot), that provides force feedback to the wrist of the surgeon with information about the force the robotic arm is experiencing, aiming to reduce the risk of damaging unexpected tissues. In this example, some factors of HRI mentioned above, like aesthetics, social context and cultural impact are of little concern. While the task being performed by the surgeon and the robot is of high importance [14].

3.1 Evaluating HRI

After designers and engineers have developed a new robot, or new software leading to different behaviour by the robot, it is important to study the performance of this interaction. When developing new robotic systems, or tuning elements of existing robots it is crucial to study the impact of the changes made as this is the only way to verify the robot's effect on the user. Some questions that are central in the analysis of HRI are the following: How does the robot's behaviour affect the user? How does the robot perceive the person who is using it? Can new users intuitively control the robot? What impact does the physical appearance of the robot have? How does a conversational robot perform in real-world social scenarios compared to a lab?

3.1.1 Approach to analyse HRI

Depending on the main task and capabilities of the robot and the goal of the research, different approaches can be taken to analyse HRI. For example, when analysing a robot safely navigating through a crowd of people, a different approach would be necessary than studying user perception of a social robot's communicative strategy. Deciding how to study and analyse such an interaction is not a straightforward task. Since HRI is a relatively new field, and it includes elements from such a broad range of robotics and interactions, there are currently no clear protocols to rate such interactions. This means that for each study, a specific set up has to be decided upon to ensure the HRI is evaluated in a proper way.

There are different ways of categorizing the ways to study HRI. One way of subdividing HRI research, as it is done in [4], [19] and [23], is by either focusing on the human perception in HRI (user oriented), the robot performance (robot oriented), or by analysing the human-robot system as a whole (system oriented).

3.1.1.1 User oriented research

When studying a robot's effect on its user and the of the participant in the study is at the center of attention, a user oriented approach is commonly used. User oriented HRI studies are those that analyze the experience of the human using the robot, tuning different attributes of the robot like behaviour, physical appearance or communicative strategies and studying its effect on the user. Often, a user centered approach is taken studying social metrics. Examples of social metrics used are: trust, empathy, anthropomorphism and engagement. Engagement is a common metric used in HRI research, as it encompasses what many studies try to achieve with their robot. As explained by [22] "engagement is the process by which two (or more) participants establish, maintain and end their perceived connection. This process includes: initial contact, negotiating a collaboration, checking that the other is still taking part in interaction, evaluating staying involved, and deciding when to end connection".

One of the advantages of the user centered approach is that the performance of the robot is not of interest. When deciding between different conversational strategies, it would be extremely time consuming to design multiple different software systems when finally only one will be implemented. This allows for a Wizard of Oz approach, a common way of doing user centered HRI studies. When participants think they are interacting with an automated robot, while in reality another human is controlling the decisions made by the robot. The Wizard of Oz approach is most commonly used in verbal social robots, but can also be used studying non-verbal behavioural attributes of the robot [4].

3.1.1.2 Robot oriented research

When trying to improve HRI as a designer or engineer, it may be more interesting to study the robot's performance in the interaction. There are many elements necessary for the robot to function properly while interacting with a human. Robot centered studies are those that evaluate the performance of the software and hardware of the robot using metrics like sensor performance, evaluation of navigation and path planning algorithms, recognition and synthesis of speech. Usually robot centered studies produce quantitative results, based on the performance of the robot. An example of this is [2], where a robot's ability to navigate through a crowd of people is being analyzed. The metrics used are all quantifiable, which means the robot's performance can be expressed in hard numbers.

An advantage of this approach is that while user centered studies require many participants to create valid results, this is often not the case for robot centered studies. This can make research easier, cheaper and more replicable. For example, when studying child-robot interaction, it might be interesting to evaluate if speech recognition software designed with adult voices also functions well with children's voices, since children's voices differ from adult voices, and often contain more incoherent utterances. For this study it might seem intuitive to use real participants, letting them speak to the robot and monitoring the performance of the speech recognition software. However, using prerecorded audio can be a lot more time efficient, while allowing easier replication of the study [12].

3.1.1.3 System oriented research

Combining robot and human centered studies is the study of the complete human-robot system. This comprises when not only the human experience or robotic performance is of interest, but the dynamic interaction or the performance of a combination of both user and robot perspective. System oriented studies are often used in medical or industrial scenarios, where the quality of the task is the ultimate objective; e.g., in surgical robotics, the quality of surgery and safety of the patient is the main goal.

Studying social robotics can also be done from a system oriented approach. One example is the study of real time user engagement analysis. This concerns the user's perspective on the robot, but also the robot's ability to detect this.

As explained in the section on user oriented research, many studies analyse user engagement after an interaction has taken place as a measure of the quality of the interaction. Recently however, the analysis of live analysis of engagement has gained popularity. Automated social robots could benefit from knowing if the conversational partner is still paying attention to the conversation, by updating its communicative strategy accordingly [17].

3.1.2 HRI data collection

Once the orientation of the study has been decided, data has to be collected to test the hypothesis. Depending on the goal of the research, different approaches to measuring HRI can be taken. Each method of collecting data has its advantages and drawbacks. Context of the acquired data is important to interpret it correctly. Many studies use more than one type of data, to cross check the results and put them into context. An example of this is [13] that tried to study an effect on user engagement elicited by the robot making eye contact with the user. The study found no increased user engagement based on the users gaze towards the robot, but the users did report higher attention towards the robot in questionnaires.

One way to subdivide the different ways to collect HRI data is in the following categories: task performance metrics, self assessment metrics, observational metrics and psycho-physiological metrics.

3.1.2.1 Task performance metrics

One of the most rudimentary ways of measuring the interaction between a human and a robot is the completion of the task being performed. This is commonly used in robot centered studies and in system centered research, where the performance of the task carried out by the system is most important [**?**].

Task performance metrics provide an objective, quantitative view of the performance of the system being evaluated, often in the form of percentages. Examples of task oriented metrics are percentage of the tasks successfully completed, time to complete the task and number of errors. Data from task performance metrics might not say anything about the user's perception of the robot. For this reason, the use of task performance metrics is not commonly used in user centered studies.

3.1.2.2 Self assessment metrics

Self assessment is amongst the most used methods of evaluation in user and system centered HRI studies. It includes surveys, questionnaires and interviews. Self assessment can be a good way of analysing the perspective of the user. As it allows the participant to elaborate on their feelings and experiences.

One of the drawbacks of self assessment is that it is hard to validate the response of the participants. Participants, especially children, may respond how they think others would respond, or how they think the researcher wants them to respond. Also, responses can be heavily influenced by the participants state of mind and disposition. Interviews can provide additional information compared to surveys and questionnaires, as they give the participants room to elaborate further on their experience. When performing an interview, it is important it is done following a predetermined structure; i.e., asking each participant the same questions in the same order, to maintain consistency [9]. During self assessment it is important to be aware of the Hawthorne effect, which refers to the tendency of participants to behave differently, knowing that they are being observed. Participants may (subconsciously) respond in favor of what is being researched [23].

3.1.2.3 Observational metrics

Another method of acquiring data for HRI evaluation is by observational analysis. Observing the HRI and counting low level micro behaviours (small, sometimes unconscious actions that can say something about a person's engagement and or liking of the robot [8]. Feelings or attention towards the interaction with the robot can be quantified through the behaviours they bring about.

Advantages of observational analysis are that these behaviours are often universal across participants regardless of age, gender or background and that observational analysis is a minimally invasive method of acquiring data.

As said in [8], behavioural cues can be subdivided into different categories; e.g., speech (including task related sentences and questions towards the robot or researcher), head movements, instances where the participant looks towards the robot or researcher, hand gestures, pointing gestures, touching of the face or body, emblems (gestures that have cultural meaning, like a thumbs up), facial expressions (instances like smiling, frowning or making a grimace) and body language (instances like leaning towards or away from the robot).

Observational analysis is commonly used in user or system centered research, where it is often used as a measure of social metrics such as engagement. An example of this is [5] which tries to use analyse user engagement in children with autism spectrum disorder. Most of the participants were non-verbal and could not read and write, so self assessment was not possible. However when a participant was actively involved in the task or conversation with the robot, this was reflected in the physical behaviours observed.

Data from observational analysis is, just like self assessment, subject to the Hawthorne effect. Therefore, only relying on observational data could give misleading results. Using self assessment as a way to establish a person's state of mind or predisposition towards the robot can be a way of overcoming this.

3.1.2.3.1 Video analysis

One of the benefits of doing observational analysis is that the interaction can be recorded to be analysed later. Most common is to have several coders watching the videos and manually scoring the behaviours. Observing in real time and keeping track of all the behaviours which are to be counted is often not feasible, so the possibility of pausing the video and re watching segments can be of help. Coding complete recordings of many participants can be very time consuming so taking randomized short sections of footage can be a good solution to keep it feasible while still gathering valid data. Using multiple human coders does require trained and independent people, and it is important strict instructions and guidelines are followed so the scoring is done consistently across different coders. [8] discusses the importance of instructing the coders, to increase inter-coder reliability.

Recent studies have tried to automate the video analysis process by using a computer to analyse video footage, or 3 dimensional body position data [1]. This greatly speeds up the process of analysing all the footage. However, a drawback of automated video analysis is that while human coders are capable of interpreting emotion and context, this is still difficult to achieve with a computer.

3.1.2.4 Psycho-physiological metrics

A more recent field of HRI analysis is that of psycho-physiological metrics. This is the analysis of physiological changes of the body as a response to behavioural stimuli. It can be used to study the direct influence of behaviour of the robot. Advantages are that participants cannot consciously influence these measurements, which can be a disadvantage in self assessment and observational analysis. Similarly, it can offer a minimally invasive way to monitor a participant's stress level and reaction to the robot. Possible measurements include the cardiovascular system (heart rate, heart rate variability, blood pressure), the respiratory system (breaths per minute), the muscular system (through electromyography, EMG) and brain activity (through electroencephalography, EEG). When using psycho-physiological data, it is important to rule out other causes for the result, like state of mind and mood; therefore, it is often combined with other methods of analysis.

An example is [15], which used heart rate, skin conductivity and EMG measurements to detect arousal in participants as a response to fast rough movements by the robot. The results show an enhanced state of arousal when the robot made unexpected movements, while this was not reported by the participants in the self assessment.

3.2 Comparison of recent studies

In the table below, 12 papers on HRI are compared. The studies have been chronologically ordered. The papers have been reviewed for the following attributes: title, aim, authors, metrics used, data acquisition, orientation, participants, used robot and limitations.

Some studies have been done to evaluate the robot they have developed, like [14] and [24] which were both designed with specific tasks in mind; while others like [18] and [1] use programmable robots like the iCub and NAO. iCub is an open source programmable robot, which was specifically designed for research on human cognition and artificial intelligence, and NAO is a commercially available robot often used in education and research.

Interesting to see is that studies that are robot oriented, often use task performance metrics to evaluate the robot performance. Sometimes a corpus of recordings is used as input in a robot oriented study, although these studies do mention the limitations of using such recordings. While it does provide a cheap and fast solution, it sometimes lacks resemblance to real world social scenarios, as the robot would have to deal with things like background noise and laughter. Studies that are user centered often use self assessment to evaluate the users experience. Often these use more participants than robot oriented studies. With the exception of [21] and [18], the relative rarity of these conditions and the large variability between participants means studying individual participants is necessary.

Title	Aim	Authors	Year	Metrics	Data type	Orientation	Participants	Robot	Limitations
Human Robot Interaction (HRI) Between a Humanoid Robot and Children with Cerebral Palsy: Experimental Framework and Measure of Engagement [18]	To evaluate if a humanoid robot can be used as an adjunct tool in rehabilitation therapy for children with CP	Malik et al.	2014	Engagement	Observational video analysis, clinical motor performance analysis	User	4 children with cerebral palsy, aged 5-14	NAO	Restless or uncooperative children.
Child Speech Recognition in Human-Robot Interaction: Evaluations and Recommendations [12]	To evaluate a number of automatic speech recognition (ASR) engines under a variety of conditions, with child speech in a variety of real-world social HRI conditions	Kennedy et al.	2017	Speech recognition	Task performance	Robot	Corpus of speech utterances of 11 children	NAO	Child grammar often differs from “normal” grammar, non-speech interpreted as speech
Evaluating the Engagement with Social Robots [1]	To evaluate the engagement aroused during interactions between humans and social robots	Anzalone et al.	2015	Engagement	Live 3D Body and head position. Gaze, joint attention, imitation	System	13 people age 20-29	iCub and NAO	No long term scenario. Only humanoids used, no androids or very human-like robots
Detection of social signals for recognizing engagement in human-robot interaction [17]	To develop systems for real time detection of engagement during a conversation	Lala et al.	2017	Engagement	Kinect, microphone. Nodding, laughter, verbal backchannels and eye gaze	Robot	Corpus of 91 conversational sessions, no number of participants mentioned	ERICA	Evaluating laughter in live scenarios is expected to be more difficult than using the corpus
Measuring engagement elicited by eye contact in Human-Robot Interaction [13]	Investigating how eye contact established by a humanoid robot affects engagement in human-robot interaction	Kompatsiari et al.	2019	Engagement	Eye tracking, self-assessment	User	24 adults	iCub	In real world scenarios, eye contact might not always indicate engagement to the same extent

Title	Aim	Authors	Year	Metrics	Data type	Orientation	Participants	Robot	Limitations
A Quantitative Technique for Analysing Robot-Human Interactions [5]	proposes a technique for quantitatively describing and analysing robot-human interactions in terms of low-level behavioural criteria	Kerstin Dautenhahn and Iain Wemy	2002	Not mentioned	Behavioural video analysis, eye gaze	User	7 children with ASD	Labo-1	Most children were non verbal, eliminating the possibility of surveys and self-assessment
The Interactive Museum Tour-Guide Robot [2]	To develop software architecture to autonomously guide visitors of a museum	Burgard et al.	1998	Navigation, mapping, path planning	Task performance	Robot	2000 visitors of a museum	RHINO	Difficulty of installation limits its use as a commercial service robot
The AESOP robot system in laparoscopic surgery, Increased risk or advantage for surgeon and patient? [14]	To examine the advantages and risks of the Automated Endoscopic System for Optical Positioning (AESOP) 3000 robot system during uncomplicated laparoscopic cholecystectomies or laparoscopic hernioplasty	Kraft et al.	2004	Voice control, manipulation	Task performance, self-assessment	System	3 surgeons, 240 patients	AESOP 3000	When more difficult extensive adhesiolysis is necessary, human assistants are still needed
Enabling Multimodal Human–Robot Interaction for the Karlsruhe Humanoid Robot [24]	Building technologies for natural multimodal human–robot interaction	Stiefelhagen et al.	2007	Person identification, sound classification, Pointing gesture detection, Dialogue processing	Task performance	Robot	15 participants	Karlsruhe Humanoid Robot	Some social cues present in HHI do not translate as well to HRI, like speech addressing

Title	Aim	Authors	Year	Metrics	Data type	Orientation	Participants	Robot	Limitations
Spoken Dialogue System for a Human-like Conversational Robot ERICA [11]	Develop a completely autonomous android that looks, behaves and interacts exactly like a human	Tatsuya Kawahara	2019	Robust automatic speech recognition, flexible dialogue, natural speech synthesis	Task performance	Robot	Corpus of 82 conversational sessions, no number of participants mentioned	ERICA	The corpus does not provide ground truth for natural social scenarios
Anxiety Detection during Human-Robot Interaction [15]	To determine the feasibility of using physiological signals to determine the human response to robot motions during direct human-robot interaction	Dana Kulić and Elizabeth Croft	2005	Arousal, motion planning	Self assessment, psychophysiological metrics: skin conductivity, EMG, heart rate	System	10 participants	CRS A460	It is still unclear whether the measured arousal indicates a measure of an involuntary reaction the subject may not be aware of or a measure of a consciously experienced emotional state such as anxiety or surprise
Empathetic Speech Synthesis and Testing for Healthcare Robots [9]	To find out if social robots with empathetic voice are acceptable for users in healthcare applications	James et al.	2020	Acceptance, empathy	Self assessment	User	120 participants in the pilot. 29 participants in the study	Healthbot	Specific applications might need voices with different emotions conveyed in them

4 Conclusion

Just as robotics has become an incredibly wide field, with implementations in almost every domain imaginable, the field of HRI has grown alongside it. Because HRI is so broad, there is no clear cut way of evaluating an interaction. When doing research on HRI, it is necessary to determine a research protocol fit for the specific goal in mind.

It is important to decide on a perspective from which to study the interaction. By either orienting on the user, robot or system as a whole. This viewpoint can help select which metrics to use, and how to collect data on these metrics.

When studying robotic performance metrics such as sensor performance, navigation, path planning algorithms, recognition and synthesis of speech the process of collecting data can be quite straightforward. As studying its performance in a task is easily quantified in numerical data. When analysing social metrics in a user or system oriented study such as trust, empathy, anthropomorphism and engagement, a different approach has to be taken to collect data. As these social metrics can not be measured directly, they have to be approximated. Common methods to do this include self assessment, observational metrics, and psycho-physiological metrics. These methods of data collection all have their advantages and disadvantages, therefore it is advised to be aware of their limitations, and combine multiple data types to validate the results and put it into context.

Psycho-physiological metrics could potentially be used more commonly in the future when less obtrusive sensors might become available, and a more direct relationship between the physiological data and psychological state is found.

References

[1] Anzalone, S. M., Boucenna, S., Ivaldi, S., and Chetouani, M. Evaluating the engagement with social robots. *International Journal of Social Robotics 7*, 4 (2015), 465–478.

[2] Burgard, W., Cremers, A., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., and Steiner, W. The interactive museum tour-guide robot. *Proceedings of AAAI-98* (01 1998), 11–18.

[3] Caprile, B., and Lazzari, G. Commanding a robot by voice: Speech and autonomous navigation for the mobile robot of maia. *Schriftenreihe der Wissenschaftlichen Landesakademie für Niederösterreich Robotics in Alpe-Adria Region* (1994), 153–157.

[4] Christoph Bartneck, Tony Belpaeme, F. E. T. K. M. K. S. *Human-robot interaction: an introduction*. Cambridge university press.

[5] Dautenhahn, K., and Werry, I. A quantitative technique for analysing robot-human interactions. 1132–1138 vol.2.

[6] Gasparetto, A., and Scalera, L. From the unimate to the delta robot: The early decades of industrial robotics. *Explorations in the History and Heritage of Machines and Mechanisms History of Mechanism and Machine Science* (2018), 284–295.

[7] Gera, D. L. *Ancient Greek ideas on speech, language, and civilization*. Oxford University Press, 2003.

[8] Giuliani, M., Mirnig, N., Stollnberger, G., Stadler, S., Buchner, R., and Tscheligi, M. Systematic analysis of video data from different human–robot interaction studies: a categorization of social signals during error situations. *Frontiers in Psychology 6* (2015).

[9] James, J., Balamurali, B. T., Watson, C. I., and Macdonald, B. Empathetic speech synthesis and testing for healthcare robots. *International Journal of Social Robotics* (2020).

[10] Karel, C. *Rossums universal robots*. Borovy, 1947.

[11] Kawahara, T. Spoken dialogue system for a human-like conversational robot erica. *Lecture Notes in Electrical Engineering 9th International Workshop on Spoken Dialogue System Technology* (2019), 65–75.

[12] Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., Senft, E., and Belpaeme, T. Child speech recognition in human-robot interaction. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (2017).

[13] Kompatsiari, K., Ciardo, F., Tommaso, D. D., and Wykowska, A. Measuring engagement elicited by eye contact in human-robot interaction. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2019).

[14] Kraft, B. M., Jager, C., Kraft, K., Leibl, B. J., and Bittner, R. The aesop robot system in laparoscopic surgery: Increased risk or advantage for surgeon and patient? *Surgical Endoscopy 18*, 8 (2004), 1216–1223.

[15] Kulic, D., and Croft, E. Anxiety detection during human-robot interaction. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2005).

[16] Kyrarini, M., Lygerakis, F., Rajavenkatanarayanan, A., Sevastopoulos, C., Nambiappan, H. R., Chaitanya, K. K., Babu, A. R., Mathew, J., and Makedon, F. A survey of robots in healthcare. *Technologies 9*, 1 (Jan 2021), 8.

[17] Lala, D., Inoue, K., Milhorat, P., and Kawahara, T. Detection of social signals for recognizing engagement in human-robot interaction.

[18] Malik, N. A., Yussof, H., Hanapiah, F. A., and Anne, S. J. Human robot interaction (hri) between a humanoid robot and children with cerebral palsy: Experimental framework and measure of engagement. *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)* (2014).

[19] Murphy, R. R., and Schreckenghost, D. Survey of metrics for human-robot interaction. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2013).

[20] Mutlu, B., and Forlizzi, J. Robots in organizations. *Proceedings of the 3rd international conference on Human robot interaction - HRI 08* (2008).

[21] Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A., and Zahari, N. I. Initial response in hri- a case study on evaluation of child with autism spectrum disorders interacting with a humanoid robot nao. *Procedia Engineering 41* (2012), 1448–1455.

[22] Sidner, C., and Dzikovska, M. Human-robot interaction: engagement between humans and robots for hosting activities. *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces* (2002).

[23] Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., and Goodrich, M. Common metrics for human-robot interaction. *Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction - HRI 06* (2006).

[24] Stiefelhagen, R., Ekenel, H. K., Fugen, C., Gieselmann, P., Holzapfel, H., Kraft, F., Nickel, K., Voit, M., and Waibel, A. Enabling multimodal human–robot interaction for the karlsruhe humanoid robot. *IEEE Transactions on Robotics 23*, 5 (2007), 840–851.

[25] TU/e. Social robotics lab: Towards humans understanding robots that understand humans, 2021.