Attitudes towards social robots and how they are affected by direct and indirect contact with a humanoid social robot

Marina Sardà Gou

Supervisors:

Professor Thomas Webb

Professor Tony Prescott

The University of Sheffield

Faculty of Science

Department of Psychology

Date of submission: 06/2020

Acknowledgements

First I would like to thank all the people who encouraged me and supported me. I would like to specially thank my supervisors Professor Thomas Webb and Professor Tony Prescott for their guidance and advise through all these years. I would also like to thank Michael Port for repairing the robots every time it was needed, and Michael Szollosy for introducing me to Pepper and NAO, and answering all my enquiries about these robots. Finally, I would also like to thank Stanislava Naneva and Daniel Camilleri for their general support.

Publications

The systematic review presented in Chapter III has been published in the *International Journal of Social Robotics*. Reference:

Naneva, S., Sarda Gou, M., Webb, T. L., & Prescott, T. (2020). A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. *International Journal of Social Robotics*, *12*(6), 1179-1201. doi: 10.1007/s12369-020-00659-4

The empirical research presented in Chapters V and VI has been accepted for publication in the research journal *Heliyon*.

Table of Contents

Extended abstract

The aim of this thesis is to investigate people's attitudes towards social robots and the factors that affect these attitudes. It uses a wide range of methodologies and aim to offer valuable and comprehensive information about how attitudes towards social robots work and how they can be manipulated. The thesis starts with a literature review which defines concepts, gives a theoretical framework, and explains the evolution of robotics and the field of Human Robot Interaction (HRI). One of the main questions that this thesis aims to answer is what are people's attitudes towards robots and what factors affect these attitudes. To address this question, a systematic review was carried out by extracting information and analysing the outcomes of studies which examined people's attitudes, acceptance, anxiety and trust towards social robots. The results of this review indicate that people typically have slightly positive attitudes towards robots, with the type of exposure to robots (e.g., direct contact, indirect contact or no contact) being the main factor affecting this outcome.

This thesis then presents a set of empirical studies that investigate whether and how attitudes towards social robots are affected by direct contact and a particular type of indirect contact (namely, extended contact). As explicit and implicit attitudes toward robots may diverge, these experiments measured both implicit and explicit attitudes. The findings suggested that direct contact affects both explicit and implicit attitudes toward social robots, while extended contact only affects implicit attitudes. Some possible explanations include supraliminal priming, a defensive reaction to a possible persuasion, or the idea that implicit attitudes are more unstable than explicit attitudes.

This thesis enables new ways of thinking and investigating human-robot relationships. Specifically, it shows that social psychology techniques (usually used to study how people interact with each other) can be applied to study how people interact with social robots. Another implication of these findings is that intergroup contact may actually affect attitudes towards robots, which could be used to create realistic opinions about social robots.

Chapter I

Introduction and overview

New advances and discoveries in robotic technologies and artificial intelligence, including robots which interact with people in social settings, have the potential to be of great use in the future of our society (Šabanović, 2010). For example, the social humanoid robot Pepper is already being used in some shops to guide and welcome customers (Pandey, & Gelin, 2018, New Atlas, 2015; Marous, 2015). Even if, in the present, the average person does not use social robots regularly, the idea of robots being present in our lives is not far away from reality (Breazeal, 2003; Broadbent, 2017). Although no one really knows exactly how this will affect our future society, many contend that robots have the potential to create a major social impact (Šabanović, 2010).

Sherry Turkle (2017) does not think that social robots will necessarily have a positive influence in society, especially on children (who will shape our future society), since they would have superficial and inauthentic interactions with robots, and misinterpret these interactions thinking that they are real connections. According to Turkle, while people interact with robots, they can potentially forget that what makes human relationships relevant is a truthful understanding of each other, something that robots may not be able to provide (Turkle, 2007). It is hard to know if bonding with robots in this way is healthy or safe. For this reason, it is necessary that both psychologists and roboticists investigate together these issues and provide answers so that people can actually benefit from the services that a robot can provide.

Peter Khan (2012) also agrees that, in the future, children will develop strong connections with social robots, the youngest being the most naïve and vulnerable. It is unknown though, how this would affect them once they become adults. An interesting point that Khan brings up is that these social robots are personified but they also allow themselves to be treated as objects. Robots do what they are programmed to do without complaining but at the same time, they elicit empathic reactions from people, partly because people are able to personify these robots. This raises the question of whether this type of relationships could actually be compared to a master-slave relationship and, if this was the case, what implications this could carry. This is why it is significant to keep on with the research done in the field of Human-Robot Interaction (HRI). These investigations could give us answers on how to design robots, how robots should behave and who should be able to interact with certain types of robots. All these with the aim of creating a society that has a healthy relationship with robotics.

Nicholas Christakis's (2019) views on the impact of robotics on society are a bit more analytical and neutral. He argues that the technology itself is not what influences people but the way it is used. A robot that behaves in a vulnerable, compassionate or altruistic way can have a positive influence on people's behaviour (Christakis, 2019; Traeger, Sebo, Jung, Scassellati, Christakis, 2020). On the other hand, a robot that acts selfishly can cause other people to behave in the same way and, therefore, may have a negative impact on society (Christakis, 2019). In other words, the right kind of robotic behaviour can actually influence positively how people behave in groups. According to Christakis, cooperation, trust, and generosity are crucial factors to have a functional society and the fact that robots may be able to modify human behaviour in such a way is highly relevant. "The aspects of AI that should concern us most are the ones that affect the core aspects of human social life—the traits that have enabled our species' survival over the millennia" (Christakis, 2019, para. 22), which are love, friendship, cooperation, and teaching (Christakis, 2019).

Unfortunately, robotics technology is advancing too rapidly for humans not to have time to evolve and adapt to this new changes (Christakis, 2019). Therefore, it is vital to know what needs to be done for robots to co-exist in our society in a constructive way. This is why it is useful to carry out studies in the field of HRI; to understand how people can benefit from robotics technologies and how robots should behave in order to have a positive impact in our world.

Because social robots are designed to work in close proximity with people who don't necessarily have expert knowledge in robotics (Alemi, Meghdari, & Ghazisaedy, 2014; Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018; Cavallo et al., 2018; Gelsomini et al., 2017; Geppert, 2004; Gombolay et al., 2018; Rantanen, Lehto, Vuorinen, & Coco, 2018), it important to understand how people will behave and react to these robots. A possible way to achieve this goal is to consider people's attitudes towards social robots. The term "attitude" does not have a widely accepted definition but different experts have proposed several ways to define this term. Allport (1935) stated that attitudes are a mental state or readiness which is affected by experience and that has an influence on people's responses to objects or situations. Eagly and Chaiken (1993) claimed that attitudes are a psychological tendency expressed by the positive or negative evaluation of a particular entity. Similarly, Eiser (1936) defined attitudes as subjective experiences involving the evaluation of something or someone. Most of these definitions coincide on the idea that attitudes have an evaluative character.

While attitudes are not the only factor affecting people's actions, they have the capability of shaping their behaviour (Ajzen, 1991; de Graaf, Allouch, & van Dijk, 2019). According to the Theory of Planned Behaviour (TPB), developed by Ajzen (1991), attitudes can change behaviour, which can be approximately predicted by observing attitudes. Therefore, investigating people's attitudes towards social robots is strongly relevant to understanding how people will behave around them.

One way of studying attitudes towards robots is by making participants interact with a robot face to face (direct contact). However, when direct contact is not always possible, certain types of indirect contact may be implemented. In social psychology, there is empirical evidence that both direct and indirect contact have a positive effect in intergroup relationships (Allport, 1954; Dovidio, Eller, & Hewstone, 2011). It could be the case that intergroup research techniques can also be applied to humanoid social robots since they can be seen as social entities (Reeves, & Nass, 1996; Reuten, van Dam, & Naber, 2018).

The principal questions that this thesis addresses are:

1. What are people's attitudes towards social robots?
2. What factors affect people's attitudes towards social robots?
3. Can techniques borrowed from research on intergroup relations be applied to HRI?
4. Does extended contact (a type of indirect contact) have an effect on people's attitudes and trust towards robots?

1.1. Overview of the approach

In order to address these questions, this thesis presents several interconnected studies that use different methodologies. First, a literature review (Chapter 2) gives the reader the understanding of relevant concepts and theories in order to provide a theoretical framework. The literature review first defines the concept of social robots as a physically embodied artificial agent that has design features which enable humans to perceive the agent as a social entity. It also provides examples of such (e.g. WABOT-1, ASIMO, Geminoid HI-4) while also explaining the evolution of these robots since the beginning of their creation. It also provides a literature review which sets the historical and contemporary background of the field of HRI. At the same time, it also describes some of the challenges faced by such a young discipline – being the lack of a unified research methodology (i.e. a unique reliable method that can be applied in many different research contexts) the primary issue.

In addition, this thesis also explores the idea that social robots can be seen as social entities or social agents, something that sets them apart from other kind of robots like, for instance, industrial robots. One of the principal theories related to this phenomenon is the so called "media equation" (Reeves, & Nass, 1996; Reuten, van Dam, & Naber, 2018), which states that people interact with communication media as if they were interacting with a real human. This is the reason why, in the field of HRI, there are a large number of studies that apply social psychology techniques in order to study how people interact with social robots.

As stated before, understanding attitudes towards robots may help to understand how people will react to these robots. With this purpose in mind, previous researchers have taken techniques from the field of psychology and adapted them in order to create methods that could quantify people's attitudes towards robots. One way to categorize attitudes is according to their class of evaluative responses, which separates attitudes into cognitive attitudes (thoughts, ideas or beliefs about a concept), affective attitudes (related to emotions, moods or feelings towards a particular concept) and behavioural attitudes (related to the individual's intentions to act in a certain way, or their actual behaviour concerning a concept) (Lemon, 1973; Eagly & Chaiken, 1993). Another system of classification separates attitudes according to their level of awareness; these are explicit attitudes (when the person having these attitudes is aware of them) (Hahn, 2014) and implicit attitudes (when the person is not aware of their attitudes) (Greenwald & Banaji, 1995). The most common methods used to measure explicit attitudes are interviews or questionnaires while, in order to measure implicit attitudes, most researchers use the Implicit

Association Test (IAT). These definitions and distinctions are expanded in detail in Chapter 2 of the thesis and they are crucial for the understanding of this work.

Apart from attitudes, this thesis also explores people's trust in social robots because, as mentioned before, social robots can be seen as social entities. Therefore, in the same way that trust between two people can help to develop a better relationship (Rempel, Holmes, & Zanna, 1985), in order to have a successful collaboration between humans and robots it is helpful to have good levels of trust in robots (Schaefer, 2013). Similar to the concept of "attitudes", the term "trust" does not have a generally accepted definition. However, all the experts agree with the idea that trust is related to the willingness of making oneself vulnerable to another person, safe in the knowledge that they will not take advantage of this vulnerability (Kassebaum, 2004; Kegan, & Rubenstein, 1973; Løgstrup, 1997; Ring, & Van de Ven, 1994). In the field of HRI, trust is also related to the expectation that a robot will perform correctly their task, without malfunctioning or causing any harm (Hancock, 2011a). Trust in robots is important because it affects how people use such technologies. A person who trusts a particular robot, will, in general, leave the robot perform its tasks autonomously and without intervening while a person who does not trust in the robot will tend to intervene more often (Hancock, 2011a). An appropriate level of trust in these robotics technologies is vital for an optimal performance. Although most methods used to asses trust in robots are qualitative or descriptive, some researchers have used quantitative self-reported methods like questionnaires (Schaefer, 2013).

Following the literature review, a systematic review is presented (Chapter 3), which aims to analyse quantitatively and rigorously integrate all of the information about people's attitudes towards social robots; therefore, providing a general view of how people perceive robots. This review illustrates these concepts in detail; it explains why they are related to people's interaction with social robots and why their understanding may help to answer the first question proposed by this thesis. This systematic review analyses different studies carried out in the field of HRI and investigates different factors that may affect people's attitudes towards robots: Namely, type of exposure to robots, domain of application, design of the robot, demographic characteristics of the participants, year of publication, and methodological quality. The outcomes of the studies were standardized and analysed statistically. The results of the review indicate that people have slightly positive attitudes towards robots, and the type of exposure to robots was identified as the principal factor affecting people's attitudes.

As stated previously, attitudes and trust can affect people's behaviour and how they interact with robots. There has, therefore, been considerable interest in understanding how attitudes develop and might be changed. Previous studies have found that having direct contact with robots can change people's attitudes and helps people to have a smother interaction with robots. However, direct contact is not always possible and this is why some studies explore some types of indirect contact like imagined or mediated contact. In social psychology, the contact hypothesis (Allport, 1954) states that contact between members of different groups (of, for example, two different ethnicities) can improve the intergroup relationship and lessen prejudice and hostility. A meta-analysis carried out by Pettigrew and Tropp (2006) which analysed 713 samples from 515 studies confirmed this theory. In fact, their results strongly suggest that intergroup contact reduces prejudice independently of the intergroup context or situation. That is to say, contact is the main variable affecting the reduction of hostility and prejudices. In addition, greater intergroup contact resulted in lower levels of prejudice.

This improvement in intergroup relationships also takes place with indirect contact (Dovidio, Eller, & Hewstone, 2011). There are three types of indirect contact; mediated contact (e.g., seeing members of the out-group in a video or the media), imagined contact (e.g., imagining an interaction with a member of the out-group), and extended contact (e.g., knowing someone who has had direct contact with a member of the out-group). Up until now, the types of indirect contact that have been investigated in HRI are imagined and mediated contact, while extended contact has been left unexplored creating an unanswered question. Therefore, one of the main aims of the empirical work in chapters 4, 5 and 6 is to investigate the effect of extended contact in attitudes and trust towards robots.

Chapter 4 is the first empirical study. There were four different conditions in this study, two of them were experimental and the other two were control conditions. In the direct contact (DC) condition, participants interacted face to face with the humanoid social robot Pepper, who acted as a hair stylist recommending hair care products to the participants. After that, they recorded two videos talking to the camera. In the first video, they explained their experience with the robot. This video was then shown to participants in the extended contact (EC) condition. In the second video, participants described an interaction with someone that they had met recently. This video was shown to participants in the extended contact control (ECC) condition. In a further control condition, participants did not interact with the robot and did not watch any video.

In order to measure attitudes and trust towards robots, this thesis used two sets of questionnaires and a task designed to measure implicit attitudes. Explicit attitudes were measured by using one of the most common questionnaires used in HRI studies, which is the Negative Attitudes towards Robots Scale (NARS), developed by Nomura (Nomura, Kanda, Suzuki, & Kato, 2004). In order to measure implicit attitudes, this study used an IAT adapted from MacDorman et al. (2009). Trust was measured by the 40 item human-robot Trust Scale, developed by Schaefer in 2013 (Schaefer, 2013). These measurements were taken before and after participants took part in the experiment.

This first study, however, provided some results that were open to multiple interpretations. One limitation was that participants in the extended contact conditions did not know the person that they were watching in the video and, therefore, it is questionable if extended contact was actually taking place (extended contact is usually defined in terms of friendships or relations with other close others). Chapter 5 therefore presents an additional empirical study, which has an improved methodology that addresses the questions presented in the previous study. The main change was made in the recruitment method. Participants had to come in pairs of friends (or someone closer than a friend like, for instance, a relative). In addition, to make sure that participants actually knew each other, they had to complete a questionnaire about their relationship. All the other details of the methodological design were the same as in the previous study. The results showed that participants in the DC condition changed positively their explicit and implicit attitudes after interacting with the robot. Participants in the EC condition changed positively their implicit attitudes after watching the video of their friend, but their explicit attitudes remained the same.

Due to the unpredicted results reported in Chapter 5, Chapter 6 presents a (pre-registered) conceptual replication to verify these findings. A different humanoid social robot which could perform in this type of experiment was selected; specifically, participants interacted with the humanoid social robot NAO in a cinema setting. The study only had two conditions; an EC condition and a control condition. Participants in the control condition did not have any intervention but, after completing all the questionnaires and tasks, they interacted with NAO and then recorded a video talking to the camera about their experience with the robot. Then, their friend, who was in the EC condition, watched the video. The findings in this study replicated the results in the previous study. Participants in the EC condition showed more positive implicit attitudes towards robots after watching the video while their explicit attitudes remained the same. The final sections of Chapter 6 explore the different possible explanations that could justify this behaviour while also discussing the limitations of this set of studies.

The last chapter (Chapter 7) summarizes the goals of the thesis and the outcomes of the studies. It also explains how the investigations carried out contribute to the field of HRI as well as the field of intergroup relations and attitudes. One of the main contributions is the information provided by the systematic review, which offers clear and concise knowledge about how people feel about social robots in a way that has not been done before. The other main contributions are the findings presented in the set of empirical studies. These findings give a good understanding of how attitudes towards social robots can be changed. Moreover, these empirical studies contribute to the field of intergroup relations since they borrow methodologies from this field and applies them to the field of HRI effectively in a concise manner, indicating that it is possible to use such methodologies with social robots. Finally, Chapter 7 also presents the limitations of the studies and gives specific ideas for future research.

This thesis positions itself in the multidisciplinary area of HRI and takes contributions from social psychology and robotics. These two apparently unrelated areas converge in the study of the relationships between humans and robots because, in order to investigate how people interact with robots, it is necessary to understand human communication and behaviour. Thus, social psychology actually plays a relevant role in HRI. Apart from this, it is obvious that having robotics knowledge is also imperative if researchers want to include real robots in their HRI studies. These are the reasons why these two areas of study (social psychology and robotics) are key in the field of HRI and why they can actually be connected.

To sum up, this thesis is composed by a literature review, a systematic review, a set of empirical studies and a final chapter discussing the outcomes these. Overall, it helps to understand people's attitudes towards social robots and the factors that can affect them. It uses innovative techniques and adapts certain methodologies used in social psychology in order to test the hypotheses. By doing this, it provides new findings which explain previous unanswered questions in the field of HRI. Lastly, it relates these outcomes to previous knowledge in the field and proposes new ideas for future research.

Chapter II

Literature Review

2.1. Overview

This chapter creates a theoretical framework while identifying previous research which has influenced the choices made in this thesis. It also provides the definitions of the main terminology used in the thesis and explains the main theories that this thesis is based upon. First, it explains why it is important to take into account the potential jobs of robotics in our society. Then, it goes on defining what a social robot is while providing some historical and contemporary background and illustrating these points with examples of real robots. Next, it defines the media equation and explains why it is a significant theory in this area. After that, it continues by describing the role of attitudes and trust in the field of Human-Robot Interaction (HRI) and explaining how these concepts can be measured empirically. Finally, it introduces the contact hypothesis theory and justifies the use of this knowledge in the HRI field.

2.2. Introduction

Social robots, which are designed to socially interact with people, will be increasingly introduced in our daily lives (Breazeal, 2003). For example, the European population is ageing since the proportion of people under 65 is decreasing while the number of those retired is expanding (Bongaarts, 2009). In this scenario, robots could be of major help in taking care of elderly and performing domestic chores (Hans, Graf, & Schraft, 2002; Harmo, Taipalus, Knuuttila, Vallet, & Halme, 2005; Kachouie, Sedighadeli, Khosla, & Chu, 2014; Roy et al., 2000; Scopelliti, Giuliani, & Fornara, 2005).

However, people do not typically live with robots and do not currently use them on a daily basis (Nomura, Kanda, & Suzuki, 2006a; Ray, Mondada, & Siegwart, 2008). If social robots are not accepted into people's homes, it is necessary to understand the main reasons which would make users take them into their environment (Venkatesh, & Morris, 2000). One of the reasons could be that robots are seen as something that could be potentially harmful both physically and psychologically (von der Pütten, & Krämer, 2015). Another issue could be the perceived incapability of the robot to perform tasks that involve social skills (von der Pütten, & Krämer, 2015). This may lead into a lack of trust towards robots and generate negative attitudes towards them.

These negative attitudes could be reinforced by the lack of experience that people have with humanoid robots. Nowadays, robots are not really present in our daily lives and interacting with them is an exceptional experience. "When deciding how to communicate with an interactional partner, people activate beliefs about the partner's abilities, knowledge, and experiences" (Kriz, Ferro, Damera, & Porter, 2010, p. 458). In other words, people rely on previous experiences when they interact with a partner. Since most people do not have experience with real humanoid robots, their knowledge likely comes from science fiction and media (Kriz et al., 2010), which may be an unrealistic reference such sources often depict robots as harmful machines or machines without social skills (Kriz et al., 2010). Because of that, some people may have unrealistic expectations or prejudices towards robots and be unwilling to interact with them. An important challenge, therefore, is to find ways to ground people's beliefs about robots in reality.

2.3. Social robots

The first idea that needs to be defined is the concept of a social robot. The present thesis defines a social robot as a physically embodied artificial agent that: a) has design features which enable humans to perceive the agent as a social entity; b) is capable of interacting with humans via a social interface (Hegel, Muhl, Wrede, Hielscher-Fastabend, & Sagerer, 2009); c) can successfully communicate verbal and/or non-verbal information to humans. In order for a social robot to be a physically embodied artificial agent, it needs to have a physical structure that mimics the behaviour, appearance, or movement of a living being (usually a human, but this could also include animals and plants). A robot can be considered to have a social interface if one of its purposes is engaging humans in social interaction. In short, a social robot is a system that can be perceived as a social entity that communicates with the user (Broekens, Heerink, & Rosendal, 2009). Social robots differ from other automated machines mostly because they are usually created to resemble a living creature (Hancock et al., 2011a).

The first robot which entered the entertainment domain and was designed to interact with people in a social context was Furby, a robotic toy designed by Tiger Electronics in 1998 (Schaefer, 2013). Furbies are pet-like robots that resemble little animals. They can learn to communicate verbally and have facial movement. Their main purpose was to interact with humans, especially children. Furbies were pioneering in introducing robots in the domestic and social context since they were the first robots that were not aimed to perform a tedious specific task. Another example of a pet-like robot would be AIBO by Sony, who was launched in 1999 (Schaefer, 2013). This robot looks and behaves like a dog. It can learn from external stimuli and is also meant to interact with people but, unlike Furby, it cannot talk.

Apart from these pet-like robots, humanoid robots have also played a crucial role in the field of social robotics. The first humanoid robot was WABOT-1, created in 1973 by Ichiro Kato. It had two arms and two legs, and it could walk like a human by controlling the electric motors (Matsusaka, 2008). Although this was not considered a social robot because it had no social interface, it was the starting point of the academic research that develops human-shaped beings by means of electromechanical technology (Matsusaka, 2008). In 2000, Honda released ASIMO, the first humanoid robot made by a commercial company. ASIMO has a social interface and, therefore, can be considered a social humanoid robot (Matsusaka, 2008).

In the following years, many other social humanoid robots were created to be used either academically or commercially. Some of these robots have achieved an impressive resemblance to the human body. For example, in 2006 Hiroshi Ishiguro's research group released the android called Geminoid HI-1. It was modelled to resemble its creator Ishiguro; it has a male body made of a metal skeleton and silicon skin which gives the robot a very realistic appearance. After that, the same laboratory created a series of other androids replicating real people such as Geminoid HI-4, Geminoid F or Otonaroid (Hiroshi Ishiguro Laboratories).

In addition to these robots, there are other humanoids that do not intend to look realistically like a human being, for instance NAO, which was Aldebaran's first robot, released in 2006. It is a humanoid of 57 centimetres with a hard shell. Its main purpose is to interact with humans and it is mainly used in universities for research purposes. Another example could be Pepper, produced by the same company, which is 120 centimetres tall and is used as a customer service robot as well as a research robot in universities.

Humanoid social robots can have many functions like assisting or entertaining people in domestic settings or recreational activities. Examples may include receptionist robots, toys, customer service robots or robotic assistants to elderly and handicapped people. Because of that, many of these robots interact, more often than not, with untrained people that have no special skills to work with a robot. Therefore, it is essential to find effective means of interaction. “The widely acknowledged shift from industrial to service robotics, and the resulting increase of robots that operate in close proximity to people, raises a number of research and design challenges” (Thrun, 2004, p. 14). The design of the interface will depend on the specific final function of the robot and developers will have to take into consideration the needs of the user. This is one of the biggest challenges that the field of robotics confronts nowadays (Thrun, 2004).

It is true that a robot that acts, talks and moves like a human will be more compelling to interact with (Kim, Park, & Sundar, 2013). Nonetheless, the Japanese roboticist Masahiro Mori (1970) observed that “as a robot’s appearance became more human-like, a robot continued to be perceived as more familiar and likeable to a viewer, until a certain point was reached (between 80% and 85% human-likeness), where the robot was regarded as more strange than familiar” (Tinwell, Grimshaw, Nabi, & Williams, 2011, pp. 743). Only when human likeliness approaches perfection and the object is almost or completely indistinguishable from a human being, familiarity rises again (Mewes & Heloir, 2009). This phenomenon is known as the "Uncanny Valley". The term "uncanny" is used to express the fact that the object creates a feeling of repulsion. Although it is just a hypothesis and it has never been validated empirically, the "Uncanny Valley" is something that could be relevant in contemporary studies because some points that researchers are dealing with are attitudes and humanoid robots, and some participants may feel uneasy interacting with a humanoid robot like Mori predicted (Mori, 1970).

Taking the social robots mentioned previously into account, it is clear that the great majority of these robots are bioinspired in one way or another since their design gives the impression that these robots could be living creatures. Although the concept of social robot is still not officially defined as such, many experts have given their point of view and provided certain traits that most social robots share, the primary trait being the ability of being perceived as a social agent that communicates with the user (Broekens, et al., 2009; Hegel, et al., 2009).

2.4. Human-Robot Interaction (HRI)

Since people can interact with communication media as if they were social agents (Nass, & Yen, 2010; Reeves, & Nass, 1996), some researchers have used psychology research techniques in order to study people's behaviour while communicating with this type of media. The discipline of Human-Computer Interaction (HCI) is a broad area of research and investigates many aspects about the interaction between people and computers such as for example, computers' usability and the design of user interfaces (Carroll, 2014). Another of these many aspects is the study of how people interact socially with computers or artificial intelligence and whether artificial agents can be seen as social agents (Hill, Ford, & Farreras, 2015; Von der Pütten, Krämer, Gratch, & Kang, 2010). It is in these kind of studies that researchers usually borrow methods previously used in psychology in order to carry out their investigations.

This last aspect of HCI is closely related to the field of HRI, which investigates the interactions between humans and robots. It is committed to recognising, developing and testing robotic systems for the human use (Goodrich, & Schultz, 2008). This particular thesis is focused on social robots, which are robots specifically designed to interact with people. Since its origins both the field of social robotics and artificial intelligence have been extremely inspired by human intelligence and appearance (Dautenhahn, 2007a). Social intelligence is a key aspect in making robots act more like humans do. However, this is a challenge that requires the cooperation of different areas of study, making the field of HRI highly multidisciplinary (Dautenhahn, 2007a). Some of the fields that study HRI include (but are not limited to) psychology, linguistics, ethics, social sciences, biology, cognitive science, robotics, mechanical engineering, artificial intelligence and computer science.

2.4.1. Historical and contemporary context of HRI

This section is aimed at providing a brief explanation of the historical and contemporary context in the field of HRI. However, it is worth noticing that this specific area of research began in the 1990s and, therefore, is still very young (Dautenhahn, 2014; Goodrich, & Schultz, 2008). The development of robotics began before that but both fields of knowledge are still new, HRI having less than 30 years of history and robotics having approximately a bit over 100 years of development. In fact, it could be said that the first robot was created by Nicola Tesla in 1898 (Goodrich, & Schultz, 2008). It was a radio-controlled boat which was, according to Tesla, “the first of a race of robots, mechanical men which will do the laborious work of the human race (O’Neill 1944, 169)” (as cited in Cheney, & Uth, 1999, p. 80). Further examples of these kind of robots include the “Electric Dog”, a tele-operated bomber created in 1923 and used in World War II (Goodrich, & Schultz, 2008). Moving forwards, apart from the robots controlled remotely, the achievements accomplished in the field of artificial intelligence began to make it possible for robots to be autonomous. The most famous early autonomous robot was called Shakey, developed from 1966 to 1972 (Nilsson, 1984). This robot was multipurpose since it was able to perceive and model its environment, perform route-finding tasks, and rearrange simple objects (Nilsson, 1984). Many experts claim that this established the basis for the growth that happened afterwards (Goodrich, & Schultz, 2008).

After that, two of the major revolutions in robotic technology were the start of behaviour-based robotics and hybrid architectures (Goodrich, & Schultz, 2008). Previously, the only robots available had a centralized system that just reacted to specific events. In other words, robots had specific reactions for each different kind of stimuli and they used a number of pre-established calculations to determine their responses. With the approach of behaviour-based robotics, robots began to have an adaptability centred systems, which meant that they could gradually adapt to their environment (Birk, 1998; Wahde, 2007). This kind of systems, which are usually bioinspired, can make mistakes at the beginning but they have a learning pattern that, later on, makes them have very coherent responses which, on top of that, are also robust to changes (Wahde, 2007).

The emergence of hybrid architectures allowed robots to have more than one system simultaneously. Regarding HRI, robots began to have both sophisticated responses (which are essential for any advanced robot nowadays) and high-level cognitive reasoning (necessary for robots to interact with humans) (Goodrich, & Schultz, 2008).“Robot behaviours initially focused on mobility, but more recent contributions seek to develop lifelike anthropomorphic behaviours, acceptable behaviours of household robots, and desirable behaviours for robots that follow, pass, or approach humans” (Goodrich, & Schultz, 2008, pp. 208-209).

HRI as a field emerged when different areas of study that were separate originally came together in order to study the interaction between humans and robots, realizing that cooperation and support across fields was essential. One of the major aims of this discipline is to make robots safe, pleasant and easy to interact with, and examine which factors affect the success or failure of such interactions. This kind of knowledge is crucial in the design of social robots but developers were already designing robots before the field of HRI even existed, hence the need of empirical data and the creation of platforms to share this specific type of knowledge such as conferences, conventions or meetings.

The first scientific meeting related to HRI was the IEEE International Symposium on Robot & Human Interactive Communication (RoMan), which took place in 1992 and it is still nowadays taking place annually around the world (Dautenhahn, 2007b; Goodrich, & Schultz, 2008; IEEE, 2010). After the creation of this symposium, there were more events that began to appear and which, similarly, were also centred in HRI. To name some of the most relevant ones, the IEEE RAS International Conference on Humanoid Robots, which was created in 2000 and focuses on the area of humanoid robots (Goodrich, & Schultz, 2008; IEEE-RAS, 2020) or the ACM/IEEE International Conference on Human-Robot Interaction since 2006 (ACM/IEEE, 2020; Dautenhahn, 2007b). Apart from that, research which is related to HRI is also presented in other kind of conferences even if they are not explicitly and specifically specialized in the field.

As previously mentioned, research about HRI began because there was a need to have empirical data in order to make better robot designs that could interact with people in a pleasant and efficient way. This knowledge can be applied to many different areas such as education, elderly care or health care, assistive robotics, entertainment, search and rescue, the military, or space exploration (Dautenhahn, 2007b; Goodrich, & Schultz, 2008; Hans, et al., 2002; Harmo, et al., 2005; Kachouie, et al., 2014; Roy et al., 2000; Scopelliti, Giuliani, & Fornara, 2005). However, this does not mean that each piece of research focuses only on one of these subjects. Even if some research is subject specific, there are a lot of other studies in the field that could be useful to many of these application areas. For example, Nomura, Kanda, Suzuki and Kato

(2004) created a scale to measure attitudes towards robots which can be used in many different areas of HRI.

Elderly and healthcare could greatly benefit from the use of social robots and, therefore, research in HRI (Schutte, 2019). For instance, robots assisting the visually challenged have to work in close proximity to people and create a sense of trust (Goodrich, & Schultz, 2008). The field of HRI could then investigate which variables affect trust in robots, and share this knowledge to those robot designers who would need it. An example of a robot used in elderly care is the robotic seal Paro, used an infinity of times in elderly care centres with great success. This robot was specifically design to assist the elderly therapeutically, and its reception has been greatly positive (Vercelli, Rainero, Ciferri, Boido, & Pirri, 2018). Another more recent example of this robots it Puffy, which is a prototype of an inflatable social robot that is meant to support children who have neurodevelopmental disorders (Gelsomini, et al., 2017). The purpose of having social robots in these setting is to provide comfort, companionship, empathy and joy while reducing anxiety, pain and distress (Dawe, Sutherland, Barco, & Broadbent, 2019). The studies investigating HRI in this kind of settings have given evidence that using social robots with elderly or patients actually improves their quality of life (Dawe et al., 2019).

Education is another area that has been of interest in the field of HRI (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018). One of the roles that have been given to robots in this area is in teaching. For example, the robot IROBI was programmed to teach English and the studies suggest that this robot improves concentration and learning compared to other kind of technologies like audio material or web applications (Belpaeme et al., 2018). Other robots act like a student’s peer, which is the case of Robovie (Ishiguro, Ono, Imai, Maeda, Kanda, & Nakatsu, 2001) or sometimes, robots act as a novice, allowing the student to act as a teacher. An example of this is the care-receiving robot (CRR), designed specifically to be a teachable robot (Tanaka, & Kimura, 2009). HRI studies focused in this area are primarily interested on the efficacy of the use of this kind of robots in the educational setting. Up until now, there is no evidence to claim that robots can teach as well as a human tutor. However, in collaboration with a human teacher, these robots support students by giving them confidence. In addition, compared to other kinds of educational technologies, educational robots provide better results in terms of learning outcomes and attention (Belpaeme et al., 2018).

Apart from that, social robots can be used as assistive robots beyond helping in the healthcare or education environment. Social service robots have great business potential to direct and track activities at restaurants, museums, and tourist centres (Pieska, Luimula, Jauhiainen, & Spiz, 2013). For example, they could be employed in a restaurant greeting customers, taking orders and carrying food to tables. As an example, Bangkok University and MK Restaurants Group Public Company Limited created several prototypes that were working in a restaurant chain in Thailand (Eksiri, & Kimura, 2015). There have been studies in the field of HRI that have focused on robots in restaurant settings. In actual fact, these robots have performed their tasks appropriately and they have been received positively (Asif, Sabeel, & Mujeeb-ur-Rahman, 2015; Eksiri, & Kimura, 2015; Pieska, et al., 2013).

Social robots have also been used for entertainment purposes, for instance, as dancers (Geppert, 2004) or dancing partners (Chen, Bhattacharjee, Beer, Ting, Hackney, Rogers, & Kemp, 2017). In the case of robots dancing on a stage, the role of the human is usually just as an observer and the interaction is minimal but, as a dancing partner, robots need to interact with humans at a close physical distance. It is the case of the robot MS DanceR (Takeda, Hirata, & Kosuge, 2007) which is one of the few robots that has been specifically design to be a dance partner. There are other robots which can be used as dancers even if they were not designed explicitly for that. For example, Chen et al. (2017) carried out a HRI study investigating the interaction between humans and robotic dancing partners. They used the robot DARCI (Dynamically Adapting Robot for Cooperative Interactions), which is multipurpose and can be programmed to dance. The results of this study suggest that people enjoyed dancing with the robot, which, in addition, helped them having some physical activity. Robotic toys can also be used by children for entertainment purposes. Some examples of these robots are Vector (Anki, 2019), WowWee CHiP (WowWee, 2020) or Ubtech Robot Alpha (Robot Advance, 2016). However, little literature has been published investigating the interaction with children and robotic toys, outside the educational or healthcare settings.

Other areas that could benefit from research in HRI are robot assisted search and rescue, space exploration, the military, or the police (Goodrich, & Schultz, 2008). It is unclear, however, if humans interact socially with the robots used in this areas. Goodrich, & Schultz (2008) provide some examples to illustrate how HRI is useful in this situations. Some of these robots are tele-operated. For example, a typical search and rescue mission could consist on using a robot to enter a dangerous area in order to search for victims of a disaster. In this case, even if the operator does not socially interact with the robot, due to the unpredictable nature of this kind of missions, the human-robot interaction in these situations is rich. Similarly, some robots used for space exploration are also tele-operated while others can be autonomous helping astronauts in exploring the surface of another planet or astronomical objects. Because of that, HRI is also aimed at improving these interactions.

Taking everything into consideration, it could be said that HRI is a growing field that has the potential of helping understand how people interact with robots. Therefore, it could help robot designers create better robots that would be easy and pleasant to interact with. Apart from that, HRI could also help understand how people perceive robots in general and what role they play in the present and future society.

2.4.2. Challenges in the field of HRI

As mentioned previously, the field of HRI is still in extremely young since it started in the 1990s (Dautenhahn, 2014; Goodrich, & Schultz, 2008). Therefore, it is encountering the challenges of such a new discipline. For example, there was not a unified measurement to evaluate people's attitudes towards robots up until 2004, when Nomura, Kanda, Suzuki and Kato (2004), created a questionnaire, named Negative Attitudes Towards Robots Scale (NARS), that was developed specifically for the purpose of having a standardized measure which could be used across different studies in the field of HRI. This gave this area of study the ability to reproduce studies by different research groups that could use the same measurements and, in this way, compare their results consistently. After that, there have been other type of scales developed to measure different aspects of the interaction between humans and robots such as the Robot Anxiety Scale (RAS) (Nomura, Suzuki, Kanda & Kato, 2006c), the Godspeed questionnaire (Bartneck, Kulić, Croft, & Zoghbi, 2009), the Almere model (Heerink, Kröse, Evers, & Wielinga, 2010) or the robot trust scale (Schaefer, 2013) all of which are explained exhaustibly in the following sections of this chapter.

Even if having these unified methods helped tremendously in the progress of HRI, it is natural in this type of multidisciplinary fields that different researchers use different techniques to investigate similar aspects of how people interact with robots (Dautenhahn, 2007b). In addition, most robots are unique designs and usually their hardware and software are not compatible with one another. Different research centres may have different kinds of robots with their distinct appearance and cognitive abilities, which also may affect the results of their studies. For this reason, it is still a challenge to design experiments using widely agreed methodologies or connecting them to previous research. "It is important to be precise about the methodological approaches used in HRI studies, but at the same time one needs to be aware that there is no ‘once-and-for-all’ solution applicable across HRI” (Dautenhahn, 2007b, p. 103).

Apart from that, another concern regarding the research methodologies in this field is the fact that, since robotics changes so rapidly, there is the risk of having a useful HRI research methodology now that may become obsolete in a few years. At the same time, results obtained in recent studies could also become outdated and non-applicable in the future. As a matter of fact, when the field of robotics was in its beginnings, the concept of “robot” was very different from what people understand as “robot” nowadays, and it will most likely keep changing as robotics keeps evolving (Dautenhahn, 2014). For these reasons and all the other explanations given previously, it is extremely problematic, if not impossible, to have a unified and timeless methodology in the study of HRI; thing that may have been affecting its thriving as a research field.

In the same line, another issue affecting experimental designs in this area is in view of the robots not being able to perform like humans during the procedure, interacting with a participant in a study. Some researchers may want to test certain variable that requires the robot to be as human as possible in terms of behaviour. The problem is that, nowadays, robotics technology does not make it possible for a robot to act like that and even if it was possible, maybe researchers would not have access to that kind of technology. Apart from that, some researchers in the field of HRI may not want to deal with the technical difficulties that suppose working with a real robot since this takes a lot of time and effort. In addition, it is sometimes the case that the use of a real robot is not necessary for certain experimental designs. In order to solve these previously mentioned issues, there have been various ways in which researchers were able to create the illusion of having a robot as capable as a human.

One of the techniques used to address this problem is the so called "Theatrical robot" (Robins, Dautenhahn, & Dubowski, 2004). This method is usually employed during the early stages of a robot design because, by the use of such, researchers can carry out their experiments without having any kind of hardware. It basically consists of using an actor who is dressed up as a robot and instructed to act like one. Typically, this person is someone with acting skills such as a mime and, in this kind of studies, this actor usually has to learn a set of scripted sentences and behaviours that then, he or she would have to perform during the procedure. Even if this method gives researchers a lot of freedom in their design, this kind of studies could be a bit problematic since they strongly rely on the abilities of the mime to actually deceive participants into believing that they are interacting with a robot.

Another method to solve the same problem is the Wizard of Oz technique (Green, Huttenrauch, & Eklundh, 2004). This method does require the use of a robot but it is not necessary for the robot to have artificial intelligence or be autonomous in any way. In this type of studies there is a person, usually a researcher that is called the "wizard", who is controlling remotely the behaviour of the robot while monitoring the interaction between the participant and the robot in question. For this method to work, deception is used; the participant should be completely naïve and they should not know that the robot is controlled by a human. The control of the robot can go from full body teleoperation to partial control. Wizard of Oz settings also give a lot of freedom to the researches in terms of experimental design and, unlike the previous method, the success of the study does not completely rely on the acting skills of the "wizard".

As mentioned previously, social robots nowadays present certain limitations regarding their social skills and interaction intelligence. So, these two techniques can be used by researchers who do not want to adapt their experimental designs according to the limitations of the robots that are available now. Furthermore, there are research centres that may also benefit from these techniques such as those who have less resources and maybe do not have access to real robots. Apart from that, it is also a fact that, for some researchers, the use of these methods can be easier that the programming of a real robot. That being said, although these techniques are still in use, as robotics technology is becoming progressively more sophisticated, there are more and more researchers that use real robots in their procedures.

HRI is a research field that is still in its beginnings and there are still many challenges that need the researchers' attention. However, as robotics technology develops, so do the research techniques used in HRI. It is clear that the multidisciplinary nature of this field allows the collaboration of experts from different backgrounds, thing that may enrich the knowledge and benefit the evolution of HRI.

2.5. The media equation

The emotions that people feel when interacting with machines which behave like living creatures could be related to the fact that people could see these machines as if they were real people, with their own thoughts and intentions (Nass, & Yen, 2010; Reeves, & Nass, 1996). In 1996, Reeves and Nass analysed this behaviour and developed the media equation: a communication theory which states that people treat computers and other new media as if they were human beings (Reeves, & Nass, 1996). The authors of this theory claim that, as a consequence, people treat media with politeness, they can feel it when machines invade their personal space, they attribute certain personality traits to media, they can also treat media as a teammate, and they also are prejudiced and assign gender stereotypes to different types of machines. Some examples of this phenomenon could be someone yelling at their TV when the reception is bad, or someone saying "Thank you" when they talk to a customer service chat bot on the phone. The media equation is also applicable to social robots (Reuten, van Dam, & Naber, 2018).

Reeves and Nass (1996) based their theory on 9 principles:

1. "Everyone responds socially and naturally to media independently of their background" (Reeves, & Nass, 1996, p. 252). The authors carried out studies with participants of all ages and professional backgrounds. All participants exhibited the media equation when using communication technologies. In other words, they treated media as if they were real people and this behaviour was not conditioned by the participants' personal differences.
2. "Media are more similar than different" (Reeves, & Nass, 1996, p. 252). Communication technologies shared common features that makes them very similar to one another. The authors of the media equation found evidence supporting their theory using many different kinds of technologies: text on a computer, a computer-controlled home theatre, small and large televisions, voices in a multimedia tutorial, and motion in political advertising. Their findings suggest that there is a remarkable similarity between these different technologies even if some of them were much more sophisticated than others.
3. "The media equation is automatic" (Reeves, & Nass, 1996, p. 252). People treat communication media as if they were humans and they do it in a natural and unconscious way. They do not need to reflect upon their reactions as they come spontaneously.
4. "Many different responses characterize the media equation" (Reeves, & Nass, 1996, p. 253). The fact that people treat communication technologies like they would treat humans can lead to many different kind of responses. These wide range of behaviours

include, but is not limited to, assigning personality, assessing the competence of the media in question, having emotional responses or having polite reactions.

5. “What seems true is more important than what is true” (Reeves, & Nass, 1996, p. 253). The authors of the media equation found out that what people perceive as being true is more important than reality itself. That is to say, the fact that a computer is perceived as being intelligent is far more relevant than the real capabilities of the computer. Perceptions are much more influential that objective facts. If people perceive that a certain technology has personality, they will respond socially whether this technology is able to have a personality or not.
6. “People respond to what is present” (Reeves, & Nass, 1996, p. 254). This principle makes reference to the fact that people react to what they observe in the media or in a technology without taking into consideration the intentions behind that media or the creators of such. For example, if someone is watching an advertisement, their reactions are related to the advertisement itself without concentrating on its persuasive intent. Similarly, when someone is playing a video game, they are interacting with the game itself instead of the programmers who developed the game.
7. “People like simplicity” (Reeves, & Nass, 1996, p. 254). The authors’ research suggests that people like media that is easy to use and understand even if this sometimes means that there is a limit to the freedom of choice. They also claim that predictable media is better received than unpredictable. When people know what to expect, they can process media in an easier way.
8. “Social and natural is easy” (Reeves, & Nass, 1996, p. 255). In line with the previous principle, which favours ease of use, the authors also claim that machines are easier to use if they follow the unwritten rules of the social and natural environments. This

is because in order to live in this world, humans have learnt how to behave following the rules for social relationships and navigating the physical world. As a consequence, these behaviours come naturally to most people and, therefore, if a machine follows these rules, it would be easier to use. This principle can be very useful for media designers or developers since they can benefit from human nature and make their products more approachable, understandable and sensible.

9. "Empirical methods show what otherwise would not be known" (Reeves, & Nass, 1996, p. 254). The authors developed the media equation by applying empirical methods used in social sciences which were not relying on introspection and, therefore, offered objective data. These techniques measured human responses and behaviour to media, which could accompany the use of focus groups or questionnaires, which are introspective methods.

The media equation has received many criticisms since it was developed. One of the main arguments that researchers have used against the media equation is the principle of anthropomorphism (Bartneck, Verbunt, Mubin, & Al Mahmud, 2007b). These researchers claim that social responses to machines can be explained by the fact that those machines are being anthropomorphised. However, Nass and Moon (2000) disagreed since they claimed that the media equation is not always explained by anthropomorphism.

Another argument against the media equation is the computer-as-medium (CAM) paradigm (Klowait, 2018). This paradigm is based on the understanding that, when people interact with a communication technology, they are actually interacting with the human behind the machine. That is to say, they are interacting with the programmers who developed an artificial intelligence or with a presenter that is talking on TV. The defenders of this paradigm also claim that this behaviour appears even if there is not a real human behind the machine. For example, if someone is interacting with a female phone chat bot, people can present gender bias while talking to the bot. This happens because the female voice could make people think that there is a female behind the programming of the bot.

In contrast with this approach, there is the computer-as-source (CAS) paradigm (Klowait, 2018). In this case, the paradigm is based on the view that people interact with the machines themselves without thinking about the presence of another human behind the machine. The media equation is a primary supporter of this approach. Taking the same example of the female phone chat bot, the CAS paradigm understands that people interact with the bot without thinking any female behind the voice. In other words, they just respond to the machine in the same way as they would with a human being and this does not necessarily mean that they anthropomorphize the bot or that they think about a real human behind the bot.

Klowait (2018) took all these last concepts into account (the media equation, anthropomorphism, CAM paradigm and CAS paradigm) and wrote a review in which he classifies three types of anthropomorphism while relating them to the media equation. The first type of anthropomorphism is the so called "anthropomorphism as human-like appearance (the 'appearance approach')" (Klowait, 2018, p. 531). This approached bases its view on the belief that humans anthropomorphise objects that have a human-like appearance. The more a machine resembles a human, the more it will be anthropomorphised. Therefore, the highest level of anthropomorphism would be achieved by a perfect android, which could not be physically distinguishable from a human being. It is unclear, though, what are the characteristics that makes a machine more or less human-like. For this reason, there will always be doubts when selecting human-like attributes for a machine, and no one can be fully sure if the right characteristics were chosen. This type of anthropomorphism can only be understood within a CAS approach. Nonetheless, it is not necessarily in line with the media equation because supporters of the media equation do not find anthropomorphism as an essential factor to make people treat machines as if they were real humans.

The second type is the “anthropomorphism as human stand-ins (the ‘stand-in approach’)” (Klowait, 2018, p. 532). This type is closely related to the CAM approach since it is based on the premise that people anthropomorphise machines when they think that there is a real human or sentient being behind the screen. The more sentient a machine appears to be, the more it will be anthropomorphised. In this case, the highest level of anthropomorphism would be a perfect artificial intelligence that would appear to have its own consciousness. In this case, a person interacting with a chat bot would attribute human intentions to the machine because they would think that either it has its own consciousness or because they would think that there is a real person behind the screen.

If this happened, then, it could be said that this machine would have passed the Turing test (Turing, 1950), which is a test that assesses an artificial intelligence’s ability to imitate human behaviour. This test was developed by Alan Turing and it has been evolved since then while maintaining its essence (Copeland, 2000; Turing, 1950; Warwick, & Shah, 2016). It is said that some artificial intelligence actually passed the test. It is the case of Eugene Goostman, a chat bot that was programmed to have the personality of a Ukrainian boy (Warwick, & Shah, 2016). However, it is arguable whether there have been some real cases in which a machine actually passed the Turing test. The experts defending this position argue that judges could have been biased favourably towards the machine. Others point out the fact that Eugene Goostman is a chat bot who is programmed to be a non-native speaker of English. This is relevant since it would have constrained the conversation and some may consider this as cheating. In any way, it was a great achievement and it is true that this kind of discussions encourage the development and perfection of new technologies. Apart from that, this provides a great example for the anthropomorphism as human stand-ins (the ‘stand-in approach’) since it looks like some judges may have thought that Eugene had a consciousness if its own.

The third and last type is the “anthropomorphism as that which makes humans interactive (the ‘interactivity approach’)” (Klowait, 2018, p. 532). This approach is extremely closely related to the media equation in the sense that it is based in the idea that people can anthropomorphise machines by interacting with them. In contrast with the other two types of anthropomorphisms, this kind of approach does not have an ideal human-like entity. It supports the idea that people can anthropomorphise objects without mistaking them for another human being. An example of this phenomenon would be a person being polite to a chat bot even when he or she knows that the bot is not human and does not anthropomorphize it.

For many media equation researchers, the term “anthropomorphism” has acquired a bad connotation. This happened because it seems that this term created an inessential association between the idea of people treating machines as if they were human beings and the idea of people attributing machines with human qualities or giving them a human-like status. That is to say, for the experts who defend the media equation, anthropomorphism is not necessary for a person to treat a machine as if it was a real person.

Taking all this into account, there is a point that has been missing in this whole discussion between defendants and detractors of the media equation. That is the fact that communication media are not real people in the same way that a humanoid robot is not a real person and, up until now, people can see that with their own eyes. That is to say, people interacting with machines usually know that they are not interacting with a human and this fact could modify their behaviour. This point could be seen as something irrelevant for some experts because sometimes people do treat machines as if they were real people. However, this would explain why some research results coming from human-human interaction studies do not exactly replicate when the same setting is applied in a human-robot interaction study. A clear example of this is the Milgram experiment developed by the psychologist Stanley Milgram (Milgram, 1963; Milgram, & Van den Haag, 1978). When this experiment has been adapted and replicated with a humanoid robot (Bartneck, Chioke, Menges, & Deckers, 2005a; Lallée, Vouloutsi, Munoz, Grechuta, Llobet, Sarda, & Verschure, 2015), the results have not been the same in the sense that, even if participants had some kind of empathy and mercy for robots, people showed more mercy for humans in the original experiment than mercy for robots in the experimental adaptations.

To sum up, maybe there are cases in which people do not treat social robots exactly as if they were other human beings. Nonetheless, this does not mean that people don’t treat robots as social agents, which this thesis defines as someone or something that can interact with people at a social level. Regarding anthropomorphism, previous research suggests that it can happen but hitherto there is no evidence that it is a necessary phenomenon for people to see machines as social agents.

2.6. Attitudes

Since social robots have the potential to be seen as social agents, some of the research carried out in the field of HRI uses techniques borrowed from psychology; and because the purpose of this thesis is to comprehend how people think about and respond to social robots, investigating people’s attitudes towards robots is one obvious way to do this. This section is aimed at giving an understanding of what attitudes are, what theories are behind the study of attitudes, and what methods have been used in order to adapt the social psychology procedures, used to investigate attitudes, to the study of HRI.

In addition, according to the Theory of Planned Behaviour (TPB), developed by Ajzen (1991), attitudes can shape behaviour and, even if attitudes and behaviour are not always necessarily and directly connected to one another, attitudes can be used to make approximate estimations on behaviour. This theory claims “that the main determinant of a behavior is a behavioral intention, which in turn is determined by attitude, subjective norms, and perceived behavioral control” (de Graaf, Allouch, & van Dijk, 2019, p. 122). Although attitudes might not be the sole factor affecting intentions and then behaviour, it is true that they can shape and influence behaviour (Steinmetz, Knappstein, Ajzen, Schmidt, & Kabst, 2016). Therefore, examining attitudes towards robots, could be useful to understand people’s behaviour and intentions towards this technology.

The study of attitudes, how they are created, how they can be measured or changed, what shapes them and how they affect behaviour has been a subject of study in social psychology since the beginning of this field (Forgas, Cooper, & Crano, 2011). The term “attitude” lacks a generally accepted definition but there are many experts on this topic that have provided several definitions to this term. Allport defined “attitude” as “a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual’s response to all objects and situations with which it is related” (Allport, 1935, p. 810). Another definition is the one provided by Eagly & Chaiken in *The Psychology of Attitudes* (1993), in which they claim that “attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavour” (p. 1). Similarly, according to Eiser (1986), an attitude is a subjective experience involving an evaluation of something or somebody. Most definitions provided by expert psychologies agree that attitudes have an evaluative character (Allport, 1935) that could be favourable or unfavourable.

The question of how attitudes are created has also been a topic of discussion and investigation in the field of psychology. Although some professionals affirm that genetics could determine an individual's attitudes, the most universally accepted assumption is that they are learnt (Lemon, 1973). Allport (1935) affirms that attitudes are created through past cumulative perceptions, feelings or experiences that are relevant to a specific topic. He also states that attitudes can be formed through the imitation of others, especially children imitating their care givers. Even if attitudes are influenced by genetic determinants, it is safe to assume that environmental factors play a crucial role in the creation of attitudes and their development.

The use of the empirical method in psychology has made it possible to design methods which can be used to measure attitudes. L. L. Thurstone (1928) and R. Likert (1932) created advanced systems to develop scales and questionnaires that made it possible to quantify attitudes. After the publication of these novel methodologies, empirical research related to attitudes increased substantially since researchers could measure attitudes in a more reliable and objective manner (Forgas, et al., 2011). Since then, there was an advancement of the understanding of attitudes.

2.6.1. The classification of attitudes

Social psychologists have classified attitudes following different criteria. In the field of HRI, it is important to understand these classifications and the characteristics of each type of attitudes because they can give HRI researchers a better perception of how people perceive robots. In the following chapters, this thesis is going to investigate (by doing a systematic review and empirical studies) these different types of attitudes towards robots. One of the oldest and a classic type of classification is the one that separates attitudes according to their class of evaluative responses (Lemon, 1973; Eagly & Chaiken,1993). According to this classification, attitudes can be cognitive, affective or behavioural (or conative). Cognitive attitudes are related to the thoughts that people have about an object or concept. Affective attitudes have to do with the feelings or emotions that arise into people's mind when they interact with or think about a concept or an object. Behavioural (or conative) attitudes involve the people's actions concerning an attitude object or a concept. Another type of classification is the one that takes into consideration the conscious availability of attitudes. They can be either explicit, when people are aware of them, or implicit, when people are not conscious of their own attitudes. Each category has its own qualities and represents different types of attitudes. For example, if a study is measuring different types of attitudes, it is not the same if participants show to have positive attitudes in one category or another. The implications of the experiment would be very different.

2.6.1.1. Explicit and implicit attitudes.

According to Hahn (Hahn, Judd, Hirsh, & Blair, 2014), **explicit attitudes** are measured by self-reported methods and necessarily involve respondents knowing that their attitudes are being assessed. Awareness seems to be of particular significance in explicit attitudes. They are usually consciously available to introspection, that is to say, people can know their explicit attitudes by thinking about them. An explicit attitude results from (a) decisions about the validity of each salient proposition as a basis for judgment and (b) attempts to maximize consistency among the different propositions (Gawronski & Bodenhausen, 2006; Gawronski, Brochu, Sritharan, & Strack, 2012). That is to say, "explicit attitudes (i.e., self-reported preferences) result from an inferential process in which a person tries to validate all of the propositions that are salient or considered relevant at the time the explicit attitude judgment is made" (Hahn et al., 2014, p. 1370).

Thurstone and Likert scales were the first formal quantitative methods to measure explicit attitudes (Thurstone, 1928; Likert, 1932). These type of techniques are aimed to measure explicit attitudes since the participant is responding a set of questions regarding his or her attitudes towards an object. Thurstone (1928) realized that, when someone expresses their attitudes verbally, the thing that researchers want to measure is not the string of words that the person has said or not even the immediate meaning of the sentence. "The opinion has interest only in so far as we interpret it as a symbol of attitude" (Thurstone, 1928, p. 531-532). He developed a sophisticated method that psychologists could use in order construct any scale to measure explicit attitudes on any topic. It is made up of statements that have a numerical value indicating how favourable or unfavourable they are to a specific topic. Participants then check the statements they agree with and then obtain a mean score indicating their attitudes.

When Likert (1932) developed his method, he made numerous references to the previous mentioned Thurstone scale and compared both methods. He pointed out several problems of the previous method although his intentions were not antagonistic. Likert proposed a new technique to create scales that is still widely used in psychology and social sciences. A Likert scale is formed by several statements which participants have to score according to their level of agreement. For example, in a study measuring attitudes towards robots, a participant may be presented the following statement: "Robots make me feel uncomfortable". Then that participant would rate that statement on a scale that usually would be from 1 to 5 (although the number of points on a Likert scale can vary depending on the study) where 1 is "Strongly disagree", 5 is "Strongly disagree" and 3 would be a neutral point. After that, a score is computed, which indicates the participant's attitudes. This method is used nowadays to develop questionnaires in an extensive range of topics in many different areas and fields. It is also the most used method to measure explicit attitudes.

In contrast to explicit attitudes, **implicit attitudes** are the ones that are consciously unavailable. They are activated automatically without the performance's awareness (Greenwald & Banaji, 1995). That is to say, the unconscious plays a big role in implicit attitudes. Thurstone (1928) already realized that participants could lie in the questionnaires. The fact that questionnaire responses could be untrue is one of the major issues in measuring attitudes using this method. "All that we can do with an attitude scale is to measure the attitude actually expressed with the full realization that the subject may be consciously hiding his true attitude or that the social pressure of the situation has made him really believe what he expresses" (Thurstone, 1928, p. 534). Participants could be deliberately lying about their attitudes or they could give answers that they consider socially desirable in fear that their response may not be well received. That is why some other methods were developed in order to measure implicit attitudes. While measuring implicit attitudes, it is more difficult for participants to lie or give dishonest responses. Therefore, they are more objective. A possible outcome when they are measured is that implicit attitudes measurement methods may reveal attitudes that could be socially unacceptable and that could make them feel uncomfortable.

Even if Thurstone had already pointed out the weaknesses of explicit measures, "[u]p until the late 1990s, research into attitudes mainly employed direct measures" (Penn, 2016, p.182). It was then when some researchers started to use more objective and implicit measures. Reaction time is one of the main measures when implicit attitudes are assessed. The most commonly used method was developed by Greenwald, McGhee and Schwartz (1998) and that is the Implicit Association Test (IAT). According to the authors, when participants are presented a task in which they have to respond (for example, pressing a button on a keyboard) to congruent associated concepts (e.g., happy + pleasant), the reaction time is faster that when the concepts are less associated (e.g., terrible + pleasant). Taking this into account, they developed a method to obtain a score from a participant's different reaction times that measure their unconscious associations.

The IAT evaluates the association between a target-concept discrimination and an attribute dimension. For example, if we wanted to measure implicit attitudes towards insects versus attitudes towards flowers, the first task would be presenting names or pictures of flowers and insects and ask participants to categorise them as such. They could do this by using a computer and pressing a key with their left hand that is associated with "insects" and another key with their right hand that is associated with "flowers". That would give the participants an idea on how to use the keyboard keys in order to perform the task. At the second step, the task would be the same as the first one but, this time, with pleasant and unpleasant words such as "happy" or "horrible", which participants would also have to categorise as "pleasant" or "unpleasant" by pressing a left or a right key on the keyboard as they did in the first task. In the third discrimination task, these two tasks would be done together at the same time. Participants would have to press the left key when they read the name of a flower or a pleasant word, and the right key when they read the name of an insect or an unpleasant word. At the fourth step of this task, the keys would be reassigned so the left key would be now for insect names and unpleasant words, and the right key for flower names and pleasant words. Finally, in the last discrimination task, the keys would be reassigned again and this time, the left key would be for names of insects and pleasant words while the right key would be for names of flowers and unpleasant words (Wittenbrink & Schwarz, 2007).

Assuming that most participants would have a more positive attitude towards flowers than insects, it would be easier for them to associate names of flowers with pleasant words, and insects with unpleasant words. Therefore, they would have a faster reaction time in those discrimination tasks. The two target concepts and the two attributes could be anything, which makes the IAT a very versatile tool to measure implicit attitudes.

2.6.1.2. Cognitive, affective and behavioural components of attitudes.

This type of classification of attitudes has a very long history that goes back to classical Greek and Hindu philosophers (McGuire, 1969, 1985). Although this is a classic division of attitudes, there is a strong connection between these three various components (Lemon, 1973).

Cognitive attitudes (also called cognitive responses to attitudes) are thoughts or ideas about the attitude object and they are often also named “beliefs” (Fishbein & Ajzen, 1975). According to Eagly & Chaiken (1993), these types of responses happen when people associate the attitude object with various attributes or when people express their attitudes verbally. These attributes can be either favourable or unfavourable and, therefore, psychologists can allocate them on an evaluative scale that can go from “extremely negative attitudes” to “extremely positive attitudes.”

A question that sometimes arises is whether attitudes that are located in a neutral point in the scale could be considered evaluative. Some psychologists prefer to categorize them as non-evaluative. For instance, if you asked a participant their attitude about a topic and they did not care about that matter, they would have attitudes located in a neutral point, which would indicate indifference. Nonetheless, neutral attitudes could also suggest that participants have evaluative attitudes which fall between positive and negative values. As reported by Eagly & Chaiken (1993), even non-evaluative beliefs express some degree of evaluation. For example, saying that someone is “active” would just state a fact but in addition it expresses a positive evaluation because the word “active” usually has positive connotations.

Affective attitudes are related to emotions, feelings or moods and they represent the amount of positive or negative feelings that people have towards an object or concept (Lemon, 1973). These affective attitudes can also range from extremely positive to extremely negative and, consequently, they can also be located on an evaluative scale (Eagly & Chaiken 1993). For example, considering the field of social robotics, when people think about humanoid robots, some of them may experience a feeling of repulsion or fear, while others may experience feelings of hope and enthusiasm. Typically, people who have favourable affective attitudes towards humanoid robots may experience positive affective reactions while people who have negative attitudes towards robots will generally experience negative affective reactions.

Behavioural (or conative) attitudes can be seen as a result of both affective and cognitive attitudes. They are related to the individual’s intentions to act in a certain way, or their actual behaviour concerning an object or a concept (Lemon, 1973). Similar to cognitive and affective attitudes, behavioural attitudes can also go from extremely positive to extremely negative; making it possible to locate them on an evaluative scale. Taking the field of social robotics into account again, some individuals may want to ban the use of service robots in certain areas while some others would like to promote their use, portraying negative and positive behavioural attitudes towards social robots respectively. It is also worth noticing that an individual does not need to act upon their attitudes in order to have behavioural attitudes. If someone has the intention to promote the robots’ use, they will have positive behavioural attitudes even if they do not carry out this intention (Eagly & Chaiken 1993).

2.6.2. Measurement of attitudes towards robots

This section presents the most commonly used scales that were specifically designed to assess attitudes towards robots. As previously mentioned, the creation of these scales provided a common method in the field of HRI that could be used across studies and, because of that, results from different research groups could be potentially compared. In the following chapters of this thesis, there will be several studies that either use or mention these scales.

2.6.2.1. The Negative Attitudes towards Robots Scale (NARS).

There are several standardized questionnaires that are meant to measure people's attitudes towards, robots. One of them is the **Negative Attitudes towards Robots Scale (NARS)**, developed in 2004 (Nomura, et al., 2004). It was developed for measuring peoples' attitudes towards communication robots in daily-life and it is based on the Likert-type scale. Its internal consistency, factorial validity and test reliability have been tested and confirmed repeatedly (Bartneck, Nomura, Kanda, Suzuki, & Kennsuke, 2005b; Bartneck, Suzuki, Kanda, & Nomura, 2007a; Cramer, Kemper, Amin, Wielinga, & Evers, 2009; Katz & Halpern, 2014; Nomura, Suzuki, Kanda & Kato, 2006b; Nomura, Shintani, Fujii, & Hokabe, 2007; Syrdal, Dautenhahn, Koay, & Walters, 2009). Table 1 shows the 14 items of the scale (1 = strongly disagree, 5 = strongly agree):

Table 1

The Negative Attitudes towards Robots Scale (NARS)

I would feel uneasy if robots really had emotions.	1	2	3	4	5
Something bad might happen if robots developed into living beings.	1	2	3	4	5
I would feel relaxed talking with robots*	1	2	3	4	5

I would feel uneasy if I was given a job where I had to use robots.	1 2 3 4 5
If robots had emotions, I would be able to make friends with them. *	1 2 3 4 5
I feel comforted being with robots that have emotions. *	1 2 3 4 5
The word “robot” means nothing to me.	1 2 3 4 5
I would feel nervous operating a robot in front of other people.	1 2 3 4 5
I would hate the idea that robots or artificial intelligences were making judgements about things.	1 2 3 4 5
I would feel very nervous just standing in front of a robot.	1 2 3 4 5
I feel that if I depend on robots too much, something bad might happen.	1 2 3 4 5
I would feel paranoid talking with a robot.	1 2 3 4 5
I am concerned that robots would be a bad influence on children.	1 2 3 4 5
I feel that in the future society will be dominated by robots.	1 2 3 4 5

*The score of these items should be reversed.

2.6.2.3. The Robot Anxiety Scale (RAS).

Another well-known questionnaire in the field of HRI, is the **Robot Anxiety Scale (RAS)**, which was also developed by Nomura, Suzuki, Kanda & Kato (2006c). This scale was designed to measure the anxiety that prevents individuals from interacting with robots having functions of communication in daily life (Nomura et al., 2006c). It is based on six–point Likert-type scale where participants had to evaluate several statements according to the level of anxiety they would feel. (1: I do not feel anxiety at all, 2: I hardly feel any anxiety, 3: I do not feel much anxiety, 4: I feel a little anxiety, 5: I feel much anxiety, 6: I feel anxiety very strongly). The RAS internal consistency, cross validity, and construct validity have been tested and confirmed repeatedly and it has also been used widely in numerous studies (Broadbent et al., 2016; Broadbent et al., 2012; de Graaf, & Allouch, 2013a; De Graaf, & Allouch, 2013b; de Graaf, Allouch, & Lutfi, 2016; Kuo et al., 2009; Nomura, Kanda, Suzuki, & Kato, 2008;

Nomura, et al., 2007; Nomura, et al., 2006c; Reich-Stiebert, Eyssel, & Hohnemann, 2019). Table 2 shows the 11 items of the scale:

Table 2

The Robot Anxiety Scale (RAS)

Robots may talk about something irrelevant during conversation	1	2	3	4	5	6
Conversation with robots may be inflexible	1	2	3	4	5	6
Robots may be unable to understand complex stories	1	2	3	4	5	6
How robots will act	1	2	3	4	5	6
What robots will do	1	2	3	4	5	6
What power robots will have	1	2	3	4	5	6
What speed robots will move at	1	2	3	4	5	6
How I should talk with robots	1	2	3	4	5	6
How I should reply to robots when they talk to me	1	2	3	4	5	6
Whether robots understand the contents of my utterance to them	1	2	3	4	5	6
I may be unable to understand the contents of robots' utterances to me	1	2	3	4	5	6

2.6.2.4. The Godspeed Questionnaire Series.

The Godspeed questionnaire does not only measure attitudes towards robots although attitudes are a big component of this questionnaire. It is extremely popular in the field of HRI, having 160 citations in Google Scholar (Weiss, & Bartneck, 2015). It was developed by Bartneck, Kulić, Croft, & Zoghbi (2009) to assess interactions with social robots. The main constructs measured in this scale are anthropomorphism, animacy, likeability, perceived intelligence and perceived safety.

Bartneck et al. (2009) explain the meaning of each construct and provide extensive explanations. Anthropomorphism makes reference to the degree in which a robot resembles the human form or behaviour. For example, an android (with flesh-like skin, realistic eyes and hair) would have a high degree of anthropomorphism while an assembly robot would present a much lower degree. In the field of robotics, it is important that robots' appearance and design match their capabilities or use since people can have different expectations depending on the robots' look. When people interact with a highly anthropomorphic, they expect the robot to behave like a human, or to be able to listen and talk. If the robot is not able to fulfil these expectations, the user may feel disappointed. In order to avoid that, robot designers should pay attention to the level of anthropomorphism of their robots (Bartneck et al. 2009).

Animacy refers to the degree in which a robot seems alive. People usually engage emotionally to robots that are lifelike (Bartneck et al., 2009). According to Bartneck et al., being alive is one of the key distinguishable standards to differentiate a robot from a human. However, because some robots have a lifelike appearance and move intentionally, people may perceive them as being alive. Sherry Turkle (1998) uses the term "sort of alive". Bartneck et al., (2009) claim that asking about a certain stimulus' presumed animacy only makes sense if it can be alive. Robots can exhibit physical behaviour, reactions to events, and even language abilities. These are usually only related to animals or humans and, consequently, one can assume that it is sensible to ask participants about their perception of the animacy of the robots.

Likeability is the construct that is strongly related to attitudes since it has to do with an evaluative response towards a robot (Bartneck et al. 2009; Eagly & Chaiken, 1993; Eiser, 1986). This evaluation can go from extremely positive to extremely negative. As stated by Bartneck et al., people evaluate other human beings within seconds of meeting a person. Since robots can be seen as social agents, they are also put in a position to receive this type of evaluations.

Perceived intelligence has to do with the degree a user perceives a robot as being intelligent. Developers of robots and artificial intelligence in general confront a big challenge in making their machines act intelligent (Bartneck et al., 2009). Many researchers in the field of HRI have been using Wizard-of-Oz settings in order to face intelligence behaviour in their robots. However, this setting can only be used in a research environment. Once the robot is used in the outside world, individuals quickly notice the robots' limitations (Bartneck et al., 2009). Apart from that, in a research environment, interactions with participants and robots usually last minutes. If users were able to interact with robots for a longer period of time, the robots' lack of intelligence would also become more apparent. Bartneck et al. (2009) also explain that evasion strategies have also been used. When communicating with the user, the robot can exhibit more or less random behaviour, and then the user may be able to see patterns in this activity which they interpret as intelligence. Nonetheless, the authors also point out that given the sufficient time, the user will realize that the robot is behaving following randomly selected patterns. After all, the perceived intelligence of a robot is strongly linked to its capabilities (Bartneck et al., 2009).

Perceived safety has to do with the degree a user perceives a robot as being safe to use (Bartneck et al., 2009). There are three different approached to make a robot safe to use: redesign the robot mechanically in order to reduce the hazard, use electronic or physical safeguards to control the hazards, and warn the user about the hazards (American National Standards Institute, 1999). Having a favourable perception of safety is a crucial requirement in order to embrace robots in human environments as partners and collaborators (Bartneck et al., 2009).

Bartneck et al. analysed all these constructs and research about the methods used to measure them. After that, they synthesised all this information and created the Godspeed questionnaire, which is based on a 5-point Liker-type scale. Table 3 shows its 23 items:

Table 3

The Godspeed questionnaire

Godspeed I: Anthropomorphism Please rate your impression of the robot on these scales:						
Fake	1	2	3	4	5	Natural
Machinelike	1	2	3	4	5	Humanlike
Unconscious	1	2	3	4	5	Conscious
Artificial	1	2	3	4	5	Lifelike
Moving rigidly	1	2	3	4	5	Moving elegantly
Godspeed II: Animacy Please rate your impression of the robot on these scales:						
Dead	1	2	3	4	5	Alive
Stagnant	1	2	3	4	5	Lively
Mechanical	1	2	3	4	5	Organic
Artificial	1	2	3	4	5	Lifelike
Inert	1	2	3	4	5	Interactive
Apathetic	1	2	3	4	5	Responsive
Godspeed III: Likeability Please rate your impression of the robot on these scales:						
Dislike	1	2	3	4	5	Like
Unfriendly	1	2	3	4	5	Friendly

Unkind	1	2	3	4	5	Kind
Unpleasant	1	2	3	4	5	Pleasant
Awful	1	2	3	4	5	Nice

Godspeed IV: Perceived intelligence

Please rate your impression of the robot on these scales:

Incompetent	1	2	3	4	5	Competent
Ignorant	1	2	3	4	5	Knowledgeable
Irresponsible	1	2	3	4	5	Responsible
Unintelligent	1	2	3	4	5	Intelligent
Foolish	1	2	3	4	5	Sensible

Godspeed V: Perceived safety

Please rate your emotional state on these scales:

Anxious	1	2	3	4	5	Relaxed
Agitated	1	2	3	4	5	Calm
Quiescent	1	2	3	4	5	Surprised

2.6.2.5. The Almere model.

There are some scales that initially were designed to assess attitudes towards technology in general but, over time, they have been evolving or being adapted into other scales that now can measure attitudes towards robot. Davis (1989) developed the initial Technology Acceptance Model (TAM), which was based on the Theory of Reasoned Action (TRA) (Fishbein, 1979; Fishbein, & Ajzen, 1975). The TAM, which describes acceptance as actual use, was used to evaluate the acceptance of many different types of technology, and the concept was modified and extended in later research. Venkatesh, Morris, Davis, and Davis (2003) presented an overview of these models and, by combining the most reliable constructs, they developed the Theory of Acceptance and Use of Technology (UTAUT) model. This scale tends to be a base for the examination of factors that determine the acceptance of social robots by older users. This is because of its repeatedly validation and the model's possible applicability to human-robot interaction.

In 2010, Heerink, et al. (2010) took inspiration from the UATUT model and adapted it into another scale called **the Almere model**. This scale was specifically designed to test the acceptance of assistive social robots by elderly users and it is based on a 5 points Likert-type scale from 1 to 5 (1 = totally disagree, 5 = totally agree). The new model has been validated using supervised experiments and observational data collected using three different social agents at older adults houses and care centres. Since its creation it has been used in several studies in order to investigate the elderly's acceptance of social robots (Louie, McColl, & Nejat, 2014; Pino, Boulay, Jouen, & Rigaud, 2015; Torta, Werner, Johnson, Juola, Cuijpers, Bazzani, & Bregman, 2014). Table 4 shows its 41 items:

Table 4

The Almere Model

If I should use the robot, I would be afraid to make mistakes with it	1	2	3	4	5
If I should use the robot, I would be afraid to break something	1	2	3	4	5
I find the robot scary	1	2	3	4	5
I find the robot intimidating	1	2	3	4	5
I think it's a good idea to use the robot	1	2	3	4	5
The robot would make life more interesting	1	2	3	4	5
It's good to make use of the robot	1	2	3	4	5

I have everything I need to use the robot	1	2	3	4	5
I know enough of the robot to make good use of it	1	2	3	4	5
I think I'll use the robot during the next few days	1	2	3	4	5
I'm certain to use the robot during the next few days	1	2	3	4	5
I plan to use the robot during the next few days	1	2	3	4	5
I think the robot can be adaptive to what I need	1	2	3	4	5
I think the robot will only do what I need at that particular moment	1	2	3	4	5
I think the robot will help me when I consider it to be necessary	1	2	3	4	5
I enjoy the robot talking to me	1	2	3	4	5
I enjoy doing things with the robot	1	2	3	4	5
I find the robot enjoyable	1	2	3	4	5
I find the robot fascinating	1	2	3	4	5
I find the robot boring	1	2	3	4	5
I think I will know quickly how to use the robot	1	2	3	4	5
I find the robot easy to use	1	2	3	4	5
I think I can use the robot without any help	1	2	3	4	5
I think I can use the robot when there is someone around to help me	1	2	3	4	5
I think I can use the robot when I have a good manual	1	2	3	4	5
I consider the robot a pleasant conversational partner	1	2	3	4	5
I find the robot pleasant to interact with	1	2	3	4	5
I feel the robot understands me	1	2	3	4	5
I think the robot is nice	1	2	3	4	5
I think the robot is useful to me	1	2	3	4	5
It would be convenient for me to have the robot	1	2	3	4	5
I think the robot can help me with many things	1	2	3	4	5
I think the staff would like me using the robot	1	2	3	4	5
I think it would give a good impression if I should use the robot	1	2	3	4	5
When interacting with the robot I felt like I'm talking to a real person	1	2	3	4	5
It sometimes felt as if the robot was really looking at me	1	2	3	4	5
I can imagine the robot to be a living creature	1	2	3	4	5
I often think the robot is not a real person	1	2	3	4	5

Sometimes the robot seems to have real feelings	1	2	3	4	5
I would trust the robot if it gave me advice	1	2	3	4	5
I would follow the advice the robot gives me	1	2	3	4	5

2.6.2.6. Implicit tasks.

The methods described above are meant to measure explicit attitudes; these are attitudes that can be expressed by the participant (Hahn et al., 2014); questionnaires are a great example since they usually ask people their opinions. Unfortunately, these answers can be susceptible to bias because sometimes people are not aware of the attitudes that are affecting their behaviour (MacDorman, Vasudevan, & Ho, 2009). Another issue is that participants may not tell the truth and it is very difficult, if not impossible, for researchers to know when a participant is giving, for example, what they believe to be a socially desirable answer in a questionnaire (MacDorman et al., 2009). For example, in the field of HRI, some participants might have more positive attitudes towards robots after they had contact with a humanoid robot just because they think that this is what it is expected from them. That is why it is useful to also measure implicit attitudes; they are unconscious and cannot be controlled by the participant (Hahn et al., 2014).

One of the most commonly used measures of implicit attitudes is the **Implicit Association Test (IAT)** (Greenwald, et al., 1998). It measures the association between two target concepts (e.g., humans and robots) and two attributes (e.g., pleasant and unpleasant). “When instructions oblige highly associated categories to share a response key, performance is faster than when less associated categories share a key” (Greenwald et al., 1998, p. 1464). That is to say, if participants have positive attitudes towards robots, then their reaction time will be shorter when they have to associate “robot” and “good”. In contrast, if they have negative attitudes towards robots, their reaction time will be longer. With this test, it is more difficult for participants to be aware of (and thus influence) their responses and as a result, researchers expect to extract data that would not be available using an explicit method like questionnaires. MacDorman et al. (2009) used the IAT in order to measure attitudes towards robots. In their research they measured implicit attitudes using as target concepts ten silhouettes of humans and ten silhouettes of robots (see Figure 1). First, they used eight pleasant words and eight unpleasant words as attributes. Then, they also used ten silhouettes of weapons and ten silhouettes of non-weapon artefacts for the attribute dimension. They obtained significant results suggesting that people have more positive attitudes towards humans rather that robots, and also that there is a stronger association between robots and weapon than there is between humans and weapons.

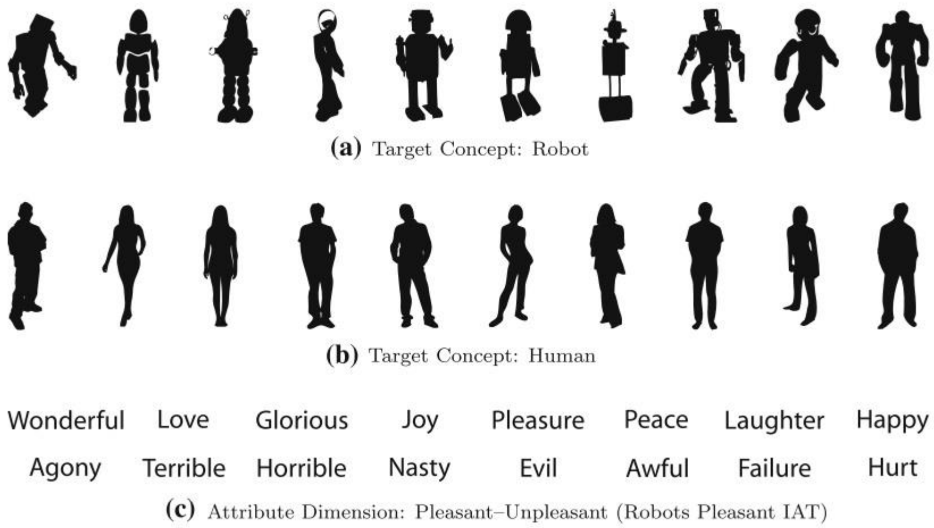


Figure 1: Images and words used by MacDorman et al. (2009) in the IAT.

2.7. Trust

The issue of trust in robots is starting to take a prominent role as social robotics is evolving towards a human-robot collaborative approach (Schaefer, 2013). The creation of appropriate levels of trust in robots is one of the most substantial obstacles to overcome in order to have successful human-robotic cooperation (Desai, Stubbs, Steinfeld, Yanco, 2009; Groom, & Nass, 2007; Schaefer, 2013). Trust in robots is key as it directly affects people's disposition to accept information provided by robots, follow robotic advice and hence benefit from the advantages that robots may provide (Freedy, DeVisser, Weltman, & Coeyman, 2007). Trust influences choices taken in unpredictable or dangerous situations (Park, Jenkins, & Jiang, 2008). For example, the more a person trusts a robot, the less they will intervene as the robot completes a task (Steinfeld, et al. 2006). Therefore, trust is crucial in maintaining pleasing relationships with robots and, for this reason, it is worthy to understand what trust is and what methods are used in order to measure trust in robots.

The concept of "trust" does not have a widely accepted definition and it has more than one connotation. According to Misztal (2013) trust is a problematic term because of its omnipresent nature. People seem to identify trust when they feel it but, having multiple dimensions, it is difficult to define. Different experts in psychology, social sciences and philosophy have provided numerous definitions. The Danish philosopher Løgstrup stated that "to trust is to lay oneself open [to the other]. […] Trust and the self-surrender that goes with it are a basic part of human life" (Løgstrup, 1997, p. 9). Ring & Van de Ven (1994) explained that there are two approaches in which one can understand trust. The first one defines trust as a risk that someone is willing to take based on their predictability of their expectations. The other approach describes trust as a belief based on the confidence in the others' goodwill.

Similarly, Kegan & Rubenstein (1973) defined trust "as a preconscious condition or attitude permitting one to enter a situation with minimal defensiveness" (p. 499). Kassebaum (2004) provides an extensive and complete definition of trust:

> Interpersonelles Vertrauen ist die auf zukünftige Ereignisse gerichtete Erwartung und das damit (in Abhängigkeit vom Ausmaß des Vertrauens und der Größe des durch ein bestimmtes Verhalten eingegangenen Risikos) einhergehende Gefühl von Ruhe und Sicherheit, dass ein oder mehrere Interaktionspartner […] ein zuvor vereinbartes, unabgesprochen wohlwollendes oder zumindest den subjektiven Erwartungen gemäßes Verhalten zeigen werden, obwohl sie die Freiheit und Möglichkeit hätten, sich anders zu verhalten, da eine Kontrolle ihrer Handlungen entweder nicht realisierbar ist oder auf diese freiwillig verzichtet wird.
>
> [Interpersonal trust is the expectation related to future actions, carried out by another person, and a sense of peace and security (depending on the degree of trust and the magnitude of the risk) that one or more interaction partners […] will act as previously agreed, in a benevolent way, or at least according to the expectations, even though they have the freedom and ability to behave differently, since control of their actions is either unrealizable or unwanted] (p. 21).

All these definitions have elements in common. First of all, they all portray trust as a mental state that an individual projects towards other people. Secondly, these definitions of trust present the concept of "vulnerability". The person who is the trustor is by definition vulnerable to the trustee, who is the person that receives the trust. That is to say, to trust someone is to place oneself in someone else's hands and make oneself vulnerable to them. And third, there has to be a certain level of expectations in order generate trust. In this scenario where the trustor is defenceless, in order for them to trust the trustee, the trustor needs to have the feeling that the trustee will not harm them. In other words, the trustor needs to expect the trustee to have good intentions. Applying this definitions taking into account social robots, one could think that trust in social robots is the willingness to make oneself vulnerable to a robot and expect the robot to behave in a harmless way without malfunctioning.

There is empirical evidence that trust influences acceptance of new technologies (Eiser, Miles, & Frewer, 2002). People determined their trust in robots by observing how robots accomplish the individual's goals and the manner in which this process is transparent (Freedy et al., 2007). Therefore, a person who trusts a robot will think that it is capable to perform its task successfully, following the instructions and without harming anyone. If that is the case, this would have a positive influence on the robot's acceptance and, consequently, a positive impact on human-robot interaction (HRI) (Hancock et al., 2011a).

2.7.1. Measurement of Trust in robots

In order to measure trust in robots, the proposed research will use the **40 item human-robot trust scale**, developed by Schaefer in 2013 (Schaefer, 2013). This scale has 40 questions and was developed to measure trust specifically in a HRI environment. It was designed to consider multiple forms of trust like cognitive trust, affective trust, and trustworthiness. Moreover, this scale can be used by any robotic domain from industry to military, to the everyday robot. This scale has been tested and validated by a set of 6 studies carried on by their developers and it has also been used in other studies since then (Kessler, Larios, Walker, Yerdon, & Hancock, 2017; Volante, Sanders, Dodge, Yerdon, & Hancock, 2016). This scale uses a percentage system score that goes from 0% (meaning that the participants has not trust in robots at all) to 100% (meaning that the participants trust robots completely). Table 5 shows its 40 items:

Table 5

The 40 item human-robot trust scale

What % of the time will this robot will...
Act consistently*
Protect people
Act as part of the team
Function successfully*
Malfunction R*
Clearly communicate
Require frequent maintenance R
Openly communicate
Have errors R *
Perform a task better than a novice human user
Know the difference between friend and foe
Provide Feedback*
Possess adequate decision-making capability
Warn people of potential risks in the environment
Meet the needs of the mission*
Provide appropriate information*
Communicate with people*
Work best with a team

Keep classified information secure

Perform exactly as instructed*

Make sensible decisions

Work in close proximity with people

Tell the truth

Perform many functions at one time

Follow directions*

What % of the time will this robot be...

Considered part of the team

Responsible

Supportive

Incompetent R

Dependable *

Friendly

Reliable *

Pleasant

Unresponsive R *

Autonomous

Predictable *

Conscious

Lifelike

A good teammate

Led astray by unexpected changes in the environment

* Items marked with an asterisk are the ones that can be used in a simplified version of the scale that only includes 14 items.

*R The score of these items should be reversed.

2.8. The Contact Hypothesis

As stated before, people do not usually interact with humanoid robots regularly and therefore, their attitudes are usually based on unreal references. Actually interacting with real robots may be one of the things that could be done to make people know more about social robots. In this way, they would know how a real robot acts and behaves. In 1954, Gordon Allport introduced the contact hypothesis in his book on *The Nature of Prejudice* (Allport, 1954). He uses the terms "in-group" members to refer to people in the same group, and "out-group" referring to people from another group. The contact hypothesis states that, under the right conditions, contact between members of different groups would improve intergroup relations and lessen hostility.

Many authors have addressed their work on the contact hypothesis. Both Allport (1954) and Saenger (1953) dedicated some chapters in their respective books to contact in intergroup relations. A shorter discussion can be found in the review by Harding, Kutner, Proshansky, and Chein (1954). Cook (1962) talked about some of the theoretical aspects of the contact hypothesis and reviewed some of the literature on this topic. References on attitude studies in general and on contact studies in particular can also be found in Williams (1947), in Arnold Rose (1947) and in Simpson and Yinger (1965).

Allport (1954) claimed that the trend of previous studies favours the conclusion that knowledge about and acquaintance with members of minority groups make for tolerant and friendly attitudes. In his work, he talks about racism and states that prejudice is reflected in both beliefs and in attitudes.

> It seems highly probable that increased knowledge of a minority group would lead directly to a truer set of beliefs. It does not follow that attitudes

> will change proportionally. One may, for example, learn that Negro blood is not different in composition from white blood without thereby learning to like Negroes. Plenty of rationalizations for prejudice are available to people who have a good deal of sound knowledge. (Allport, 1954, p. 255)

That is to say, the fact that one has their beliefs based on realistic references does not necessarily mean that their prejudiced views will change automatically. In spite of this, Allport also affirms that intergroup tension, hostility and prejudice can be reduced by contact between groups of people; especially contact that brings knowledge and acquaintance.

However, one may also think that the nature of the interaction could be relevant in these studies. In fact, Cook and Selltiz (1955) argued that, while most studies reported that contact lead to favourable changes between groups, there were other researchers who had found that contact resulted in favourable changes on the part of some participants, in no change on the part of others, and in unfavourable changes for others. For this reason, it is important to be careful with indiscriminative generalizations that could be misleading and be cautious extracting conclusions from the available evidence.

In the past years, the contact hypothesis has been extensively tested and different types of intergroup contact (direct contact, extended contact, mediated contact and imagined contact) have been researched (Amir, 1969; Crisp & Turner, 2009; Dovidio, Eller, & Hewstone, 2011; Ortiz & Harwood, 2007; Pettigrew, Christ, Wagner, & Stellmacher, 2007; Schiappa, Gregg, & Hewes, 2005; Wright, Aron, McLaughlin-Volpe, & Ropp, 1997).

2.8.1. Direct contact

Empirical work suggests that **direct contact** can affect peoples' attitudes towards members of minority groups (Amir, 1969; Dovidio et al., 2011; Pettigrew et al., 2007). Direct contact was the first to be tested on the field of contact studies. As stated before, most studies in this field have reported a decrease in hostility between groups as a result of a direct contact situation.

Following this line of research, there have been some studies which investigated direct contact with humanoid robots and its effects on attitudes towards them. In these studies, participants interact directly with a robot either performing a task or talking with the robot. The empirical evidence indicates that repeated interaction with robots can change attitudes towards them. They can be more positive or negative depending on the nature of the interaction; a positive interaction may lead to more positive attitudes and a negative interaction may cause more negative attitudes towards robots (Nomura, et al., 2008; Nomura, Kanda, Yamada, & Suzuki, 2011).

2.8.2. Indirect Contact

Nevertheless, direct contact has its disadvantages. Sometimes direct contact can be difficult to be implemented successfully because of the given anxiety and hostility that sometimes exists in intergroup relations. This anxiety and hostility may cause a negative outcome from a direct contact interaction (Ortiz & Harwood, 2007). In the field of Human-Robot Interaction, the main disadvantage is the fact that direct contact is not always possible because some research facilities lack the necessary resources to run an experiment with an actual humanoid robot. Moreover, since people do not live together with robots (Nomura et al., 2006b; Ray, Mondada, & Siegwart, 2008), they do not have direct contact regularly. In these cases, researchers implement indirect contact and the participants do not interact with a tangible robot.

An example of indirect contact is **mediated contact**, which involves contact via media like television or videotapes. According to Bandura (2009), humans have the capacity to learn through observation and experiences can be gained both through direct contact and through models observed in the media. Ortiz and Harwood (2007), claimed that audience members can learn positive intergroup behaviours by watching televised representations of characters engaging in favourable intergroup interactions. So, exposure to positive intergroup contact on television would be associated with more positive intergroup attitudes.

Indirect contact also includes **extended contact**: “learning that an in-group member is friends with an out-group member” (Dovidio et al., 2011, p. 147). This concept was first proposed by Wright et al. (1997); they suggested that knowing that an in-group member has a close relationship with an out-group member can lead to more positive attitudes and less preconceived ideas about out-group members. Wright et al. claim that in-group members that are friends with members of the out-group provide understanding and are a source of knowledge about the out-group. This encourages tolerance and acceptance between groups. In their research, Wright et al. (1997) provide direct causal evidence for the extended contact hypothesis.

It is shown in previous studies that a research method used only in humans can be exported and adapted in order to make research between humans and robots. That is to say, humans’ behaviours towards other humans are comparable to humans’ behaviours towards robots (Kuchenbrandt & Eyssel, 2012). Since indirect contact can change intergroup attitudes, it could be hypothesised that indirect contact could influence attitudes towards robots too.

Another instance of indirect contact is **imagined contact**, which is imagining an interaction with an out-group member. This technique was first used by Richard J. Crisp in 2009. He was investigating about the interaction between members of different groups. He stated that "encouraging people to mentally simulate a positive intergroup encounter leads to improved out-group attitudes and reduced stereotyping." (Crisp & Turner, 2009, p. 231). The main idea here is that imagining an interaction with an out-group member activates processes in the mind that are parallel to those involved in actual contact (Crisp & Turner, 2012). Therefore, imagined contact could be used as a first step before direct contact in order to improve intergroup relationships. This technique has also been used in Human-Robot Interaction (HRI) and it was shown that it improved participants' attitudes towards robots. "After imagining contact with the robot [NAO], participants indicated less negative attitudes and less anxiety towards robots." (Kuchenbrandt & Eyssel, 2012, p. 463)

One of the main differences between different types of contact is the way researchers can control and monitor the stimuli. Extended contact can be planned and monitored to a greater degree than imagined contact. This type of contact is usually implemented by telling the participant to imagine a situation in which they interact with an out-group member (Crisp & Turner, 2009). In this case, the stimulus is what the participant is imagining. Researchers cannot see what a participant has in their mind and, therefore, it is more difficult to control and monitor how imagined contact is affecting the participant’s attitudes. With extended contact, researchers have a greater degree of control over the intervention since, in this case, the stimulus is the knowledge that an in-group member has met or has a relationship with an out-group member. The same happens with direct contact, which can be easily monitored by recording the in-group member interacting with the out-group member, and mediated contact, which can also be easily monitored by keeping the video used to implement the contact.

In addition, extended contact is the only type of contact that involves two people in the in-group; the person who knows the out-group member (who creates the extended contact) and the person who knows about this intergroup relationship (who receives the extended contact). This means that this type of contact can use the mechanisms of in-group norms. That is to say, extended contact is focused on another in-group member and their perception about an out-group member. If this in-group member has a relationship with an out-group member, this could create a new in-group behavioural norm implying that out-group members are not threatening. This happens because usually, in-group members can identify themselves easily with other in-group members.

Apart from that, this in-group member (who is providing the extended contact) already knows the cultural background and behavioural customs that compose their in-group and, therefore, they could implement extended contact in a more relatable way. As an example, we could take the case of two groups of people who speak different languages (for instance, Chinese and Spanish) and have to coexist in the same environment. If they were to have direct or mediated contact, the intervention may not work in a fluent way because of a language barrier. If they had imagined contact, they would probably imagine the out-group member talking in their own in-group native language, which is completely unrealistic. However, if they wanted to implement extended contact, they would need only one person in each group who spoke fluently the out-group language. One member of the Chinese community could meet a member of the Spanish community and then explain their experience to their respective in-group members. In this way, extended contact would be provided by an in-group member using the in-group native language and giving a particular view from an in-group perspective that could be easily understood by the in-group members. These members would be able see how their peer reacts to the out-group and, therefore, this experience has the potential to be more relatable.

In the case of contact with robots, one could argue that extended contact is more difficult to implement than imagined contact because extended contact requires someone interacting with a robot at least once. In contrast, imagined contact does not need this previous interaction with a robot, and therefore it does not require this type of equipment, making it more affordable and easier. However, one of the advantages of extended contact is that it is able to provide a more realistic reference than imagined contact. This is because imagined contact relies on previous experiences with the out-group. Since most people do not usually interact with social robots, it would be difficult for them to imagine themselves interacting with one. Some previous studies have shown a picture of the robot NAO to participants before they were involved with imagined contact with this robot (Wullenkord & Eyssel, 2014). However, it is uncertain how the fact that they actually saw the picture of the robot before the intervention affected their experience with imagined contact and, therefore, the results. In this sense, extended contact would provide a more realistic reference since, at some point, an in-group member would need to have an interaction with a real robot and, therefore, their reference would be based on an actual robot.

For the empirical work in the present thesis, extended contact was implemented in a HRI setting. There were several factors that contributed to this decision. The first and most important is the fact that extended contact with robots has never been implemented before in an empirical study. Moreover, extended contact causes less anxiety than direct contact, it instigates in-group norms and it is based on a realistic reference.

To sum up, there is empirical evidence that both direct and some indirect contact can improve intergroup relationships in communities with different ethnical groups. Intergroup methodologies have been adapted successfully in the field of HRI because methodological techniques that are usually applied to social psychology can also be applied to study relationships between robots and humans.

2.9. Overview of the present thesis

The present thesis will address the following questions. First of all, it will investigate attitudes towards robots by carrying out a systematic review that will analyse the empirical evidence that, up until now, have performed experiments on attitudes, anxiety, acceptance and trust towards social robots. The results of these studies will be standardised and categorized, and there will be a comparative analysis in order to find out what are the factors that affect the most how people perceive robots. After that, this thesis will complete a series of empirical studies investigating the effect of direct and extended contact on attitudes and trust towards robots, thing that has never been investigated before. These studies will have both implicit and explicit measures since this will provide a more complete view on the participants' perceptions of robots.

Chapter III

Study 1: A systematic review of attitudes, anxiety, acceptance, and trust towards social robots

3.1. Overview

This chapter presents a systematic review which has analysed the empirical research carried out to investigate attitudes, anxiety, acceptance, and trust towards social robots. This study was carried out in collaboration with another Ph.D. student, Stanislava Naneva, who is also in the field of HRI.

First, I completed a literature search following a previously established protocol. After that, all the collected studies went under a process of selection also following previously established inclusion and exclusion criteria. After that, Stanislava Naneva, extracted the data from the studies also following a pre-established protocol and, in this case, I was the second coder. Then, she performed the analysis that was discussed among us and our supervisors.

This systematic review has been published in the *International Journal of Social Robotics*. Reference:

Naneva, S., Sarda Gou, M., Webb, T. L., & Prescott, T. (2020). A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. *International Journal of Social Robotics*, *12*(6), 1179-1201. doi: 10.1007/s12369-020-00659-4

3.2. Introduction

According to a widely-reported large-scale survey (European Commission, 2012), a substantial proportion of EU citizens have negative attitudes towards the use of robots within healthcare and other fields that are traditionally dominated by humans. There have also been suggestions of a growing public anxiety that automation, enabled by robotics, will lead to a significant loss of jobs (Broadbent et al., 2012; Ebel, 1986). As we will explore in this chapter, attitudes towards the use of robots appear mixed, dependant on the setting and question asked, and in some cases somewhat divorced from reality (e.g., there is evidence that attitudes are based on science-fiction, rather than objective reality; (Kriz, et al., 2010). While attitudes do not consistently predict behaviour, they are thought to influence people's behavioural intentions (Ajzen,1991) and therefore may predict the uptake and use of robots alongside other variables such as anxiety, trust, and intention to use and engage with robots. An improved understanding of people’s attitudes towards robots should therefore help to inform future research, development, and deployment of robots in various domains of public and private life.

The present review focuses on social robots, due to their increasing use in various settings such as healthcare, entertainment, and customer service (Pieska et al., 2013; Takeda et al, 2007; Hancock et al., 2011a). While the idea of robots that can interact socially with people has been around for some time, their use has been relatively limited and less widespread in comparison to, for example, manufacturing robots (Ray et al., 2008; Nomura, et al., 2006a) Nevertheless, social robots garner attention from the media and general public alike, and have sparked debate about their potential impact on society (Nørskov, 2017; Zhao, & Yi, 2006). We define a social robot as a physically embodied artificial agent (i.e., something that has a physical structure that mimics the behaviour, appearance, or movement of a living being - usually a human, but could also be an animal or plant) that: (a) has features that enable humans to perceive the agent as a social entity (e.g., eyes); (b) is capable of interacting with humans via a social interface (Hegel, et al., 2009); and (c) can communicate verbal and/or non-verbal information to humans (see Supplementary Materials 2). In short, a social robot is an embodied system that can be perceived as a social entity and is capable of communicating with the user (Broekens, et al., 2009).

To date, no systematic review has investigated and synthesised the current evidence on people's attitudes toward, trust in, anxiety associated with, and acceptance of social robots. Evidence suggest that all of these beliefs can predict the use of social robots (Heerink et al., 2010; Im, Hong, & Kang, 2011), and reflect the same broad construct (Gaudiello, Zibetti, Lefort, Chetouani, & Ivaldi, 2016; Gombolay et al., 2018; Herse et al., 2018; Li, Rau, & Li, 2010), which is people's perception or evaluation of robots.

3.2.1. Attitudes towards social robots

Current evidence on people's attitudes toward social robots reveals a somewhat ambiguous picture that makes it difficult to say whether people, in general, have a negative or positive view of social robots. This is, at least to some extent, likely to be due to the variety of contexts in which social robots are employed. People generally agree that, while working alongside robots is not out of the question, robots should not entirely replace humans in jobs that require substantial social skills (e.g., nursing; (Enz, Diruf, Spielhagen, Zoll, & Vargas, 2011)). At the same time, some studies have found positive attitudes toward robots performing jobs that demand more social skills (European Commission, 2012; Enz et al., 2011). These inconsistencies merit further investigation.

In addition to providing an overall assessment of the current evidence of people's attitudes toward robots, where possible, the present review will also look at three distinct components of attitude – cognition, affect, and behaviour (Breckler, 1984). Cognitive attitudes reflect people's thoughts – or cognitive evaluations - about the attitude object (e.g., that robots are useful). Affective attitudes reflect the individual's feelings or emotions toward the attitude object (e.g., whether they feel warm toward social robots). Finally, behavioural attitudes reflect people's observable or self-reported behaviours toward an attitude object (e.g., the extent to which they approach and interact with a social robot). Distinguishing between the various components of attitude may provide more insight into people's attitudes toward social robots, and potentially account for some of the mixed findings identified in the literature to date (e.g., people may have positive cognitive attitudes, believing that social robots are worthwhile, but have negative affective attitudes, to the extent that they feel uneasy when they think about interacting with a robot).

3.2.2. Anxiety about social robots

A number of studies provide evidence that anxiety, alongside attitudes, predicts intentions to use social robots and the quality of people's interaction with social robots (Nomura et al., 2006b; Nomura et al., 2008; Nomura et al., 2011). Anxiety toward robots is often measured using self-report measures, such as the Robot Anxiety Scale (RAS; (Nomura et al., 2006b)) or direct observation of behaviour during human-robot interaction (HRI). Despite the potential importance of anxiety in shaping how people interact with robots, current evidence presents a mixed picture as to how anxious people are about social robots. For example, Nomura, Shintani, Fujii, and Hokabe (2007) found that both anxiety and attitudes can affect how people behave during HRI in similar ways, while de Graaf and Allouch (2013) found that participants interacting with a robot showed a change in their anxiety but not their attitudes. Therefore, the present review sought to integrate the evidence on anxiety to date, as well as identify factors that might account for the variable estimates in individual studies.

3.2.3. Trust in social robots

Similar to anxiety, trust has been recognised as a factor that, at least in part, predicts not only the quality of HRI but also how willing people are to use social robots for certain tasks (Salem, Lakatos, Amirabdollahian, & Dautenhahn, 2015). Trust is likely to be particularly important in relation to social robots, especially in healthcare, where it has been associated with patient satisfaction and therapeutic effectiveness (Hall, Dugan, Zheng, & Mishra, 2001). So far, reviews have focused on the impact of trust in robots on human-robot interaction, showing that the main factors influencing trust relate to aspects of the robot (e.g., the robot’s design and performance) while environmental factors play a more moderate role in how much people trust robots (Hancock et al., 2011a). However, the impact of trust in relation to social robots specifically has not been reviewed (Hancock, Billings, & Schaefer, 2011b).

3.2.4. Acceptance of social robots

Acceptance is generally defined as the intention to use, and in some cases, as the actual use of robots (Davis, 1989; Heerink, et al., 2010; Venkatesh, et al., 2003). Compared to anxiety and trust, there is considerably more evidence on the extent to which people accept social robots, particularly in the healthcare and elderly care domains. Acceptance of robots in healthcare has been found to be mixed and can vary considerably depending on the function and appearance of the robot (Broadbent, Stafford, & MacDonald, 2009). Despite the potential that social robots have to alleviate the ever-growing demands on healthcare professionals (Broadbent et al., 2009; Dawe, et al., 2019), low levels of acceptance can prove detrimental to the development and utilisation of such technology (Broadbent et al., 2009; Klamer, & Allouch, 2010). Therefore, a broader understanding of the extent to which social robots are accepted in healthcare and other settings; along with factors that are associated with acceptance is needed.

3.3. What factors influence people's attitudes towards robots?

Several factors are likely to be associated with people's attitudes towards, trust in, acceptance of, and anxiety towards social robots. For example, people's beliefs may differ as a function of whether they have recently been exposed to social robots (e.g., studies that provide direct HRI may report different attitudes to studies where participants do not interact with a robot), the intended domain of application (e.g., companionship and domestic assistance, education, healthcare), and the design of the robot (e.g., humanoid, anthropomorphic). We expand on these potential factors below.

3.3.1. Type of exposure to robots

The way that people think about robots might be affected by whether they are given the opportunity to interact with a robot, directly or indirectly prior to their attitudes being measured. Studies generally provide participants with at least one of three types of exposure to robots (i.e., HRI) which we explore in this review:

No HRI - participants were not asked to interact, view, or imagine a social robot or robots (e.g., participants were only asked about their attitudes towards social robots in general (de Graaf, Allouch, & Lutfi, 2016));

Indirect HRI - participants observed a direct interaction or were shown (or asked to imagine) a representation of the social robot or robots (e.g., participants read an illustrated description of a NAO robot; (Reich-Stiebert et al., 2019));

Direct HRI - participants interacted with a social robot that was physically present at the same time and place as them (e.g., participants took part in a mock-interview with a Geminoid HI-2 robot (Zlotowski, Sumioka, Nishio, Glas, Bartneck, & Ishiguro, 2015)).

3.3.2. Domain of application

Evidence suggests that people's attitudes toward robots may, to some extent, depend on the domain in which the robot is (or is intended to be) used (May, Holler, Bethel, Strawderman, Carruth, & Usher, 2017; Savela, Turja, & Oksanen, 2018). For the purposes of this review, we identified six broad domains of application:

Companion robotics and domestic assistance - robots designed specifically and exclusively to interact socially with humans for a prolonged period of time and to provide companionship (e.g., a study investigates attitudes towards the robots NAO and Darwin depending on its appearance and facial expression using a scale that measures trust and affective attitudes (Hosseini et al., 2017)); or robots that are designed to help with domestic chores as well as provide social interaction (e.g., a study investigating the evaluation of a socially assistive robot in a smart home setting; (Torta et al., 2014));

Education - robots designed to assist educators with teaching and social interaction with students (e.g., a study investigating how students evaluate the use of NAO to teach English lessons; (Alemi, Meghdari, & Ghazisaedy, 2014)).

Healthcare - robots designed to help patients, doctors or healthcare providers (e.g., a study investigating the attitudes and preferences of staff, residents and relatives of residents in a retirement village towards a health-care robot; (Broadbent et al., 2012)).

Paediatric care - robots that are used in healthcare but specifically designed to assist children and the healthcare providers who treat them (e.g., an evaluation of physiotherapists' acceptance of assistive robots as a therapeutic aid for children in rehabilitation; (Carrillo, Butchart, Kruse, Scheinberg, Wise, & McCarthy, 2018)).

HRI - robots that are designed primarily to interact with people, with any additional functionality (e.g., providing care) being secondary. For example, playing games or having a conversation (e.g., a study examining the effect of group size on people's attitudes and behaviours toward robots as interaction partners; (Chang, White, Park, Holm, & Šabanović, 2012)).

General application - the study does not specify or imply an application domain for the robot or robots being investigated. (e.g., a study investigating the effectiveness of exhibitions of robots as a means of shaping people's beliefs about robots; (Kim, Lee, Aichi, Morishita, & Makino, 2016)).

3.3.3. Design of robot

Design features of robots, such as the degree of human-likeness, are likely to influence people's attitudes towards robots (Hancock, et al., 2011b; de Graaf, & Allouch, 2013); however, this influence has not been quantified or reviewed comprehensively so far. The present review therefore categorised each of the robots studied into one of three broad categories:

Humanoid - a robot that resembles a human body (e.g., the humanoid robot NAO; (Serholt, Basedow, Barendregt, & Obaid, 2014)).

Anthropomorphic - a robot that imitates some parts of the human body and can be subject to anthropomorphism by the user (e.g., a robot with a human-like face; (Dunst, Trivette, Prior, Hamby, & Embler, 2013)).

Non-humanoid - a robot that resembles any other living organism except for a human or does not imitate a living organism (e.g., Aibo, a robot that resembles a dog; (Bartneck, et al., 2007a))

3.3.4. Geographical location

The cultural background and nationality of users may contribute to the variability in people's attitudes toward (Bartneck, Nomura, Kanda, Suzuki & Kato, 2005), trust in (Li, Rau & Li, 2010), and acceptance of (Bernotat & Eyssel, 2018) social robots. The present review therefore compares the geographical locations (i.e., countries) in which the studies took place as an approximation of participants' cultural backgrounds. Enough data was available to compare eight geographical locations: Australia, France, Germany, Italy, Japan, the Netherlands, New Zealand, South Korea, Taiwan, and the United States of America (USA).

3.3.5. Sample characteristics

Attitudes towards robots also likely vary according to demographic factors such as users' age and gender (de Graaf, & Allouch, 2013). For example, in general men tend to have more positive attitudes towards robots than women (May et al., 2017). Similarly, young adults tend to have more positive attitudes towards robots than elderly adults and are more willing to make use of robots (May et al., 2017). Therefore, the present review investigates whether participants age and gender are associated with their beliefs about robots. In addition, some studies have reported that previous experience with and long-term exposure to robots also affects people's attitudes (Leite, Martinho, & Paiva, 2013) which is why the present review also attempted to investigate this factor.

3.4. The present review

The present review expands on earlier efforts to understand people's beliefs about social robots (e.g. Hancock, et al, 2011b; Broadbent et al., 2009; Savela et al., 2018; Chen, & Chan, 2011) by taking a broad approach to the collection and synthesis of available literature in order to provide an overview, of not only people's attitudes toward social robots, but also other beliefs which are relevant to the uptake of robotics such as acceptance, anxiety and trust. The review sought to include studies focusing on any type of social robot and a wide variety of domains where they might be used. In addition, we also present a series of analyses that go beyond previous systematic reviews. Specifically, we have developed a novel method for standardising the measures of participants' beliefs about robots in each of the primary studies. This approach enabled us to estimate people's attitudes toward robots, across the available evidence, weighing each estimate by the size of the sample in a manner similar, but not identical, to that of a conventional meta-analysis. Additionally, by combining estimates of beliefs in specific areas (e.g., studies focusing on social robots in particular contexts), we were able to investigate the factors that are associated with people's attitudes toward robots.

3.5. Method

This review was pre-registered on PROSPERO (CRD42017057331).

3.5.1. Systematic literature search

In order to identify studies that measured people's attitudes toward, trust in, acceptance of, and / or anxiety toward social robots, the following databases were searched between January and February, 2018 and repeatedly searched in January 2019: PsycINFO and PsycARTICLES (Ovid), IEEE Xplore, ProQuest, and Google Scholar. A separate search was conducted for each of the four measures of interest in each database (except Google Scholar) using the search terms: "[attitud* / accept* / trust* / anxi*] AND (robot* OR "human-robot interaction" OR "assistive robot" OR "social robot") AND participant". A slightly different approach was used for Google Scholar as it was found that the combination of the above search terms did not generate as relevant results as the phrase: "[attitude / acceptance / trust / anxiety] AND robot AND participant". Only articles from the first ten pages of results for each of the four searches conducted in Google scholar were considered in order to ensure that the search was manageable. In order to identify further grey literature, publication lists of relevant research laboratories were also searched (a full list of the laboratories can be found in the review's protocol on PROSPERO). No limitations on publication date were specified for any of the databases. The references of the identified papers were added and managed via EndNote where duplicates were removed prior to screening the research articles. Figure 2 shows the number of articles that were identified as well as the number of articles that were included and excluded at each stage of the screening process.

3.5.2. Screening and selection of relevant papers

The search results were screened by a member of the research team in two stages and guided by a priori inclusion and exclusion criteria. Any uncertainty as to whether a paper should be included or not was resolved through discussion with the research team.

First, the titles and abstracts of the retrieved research articles were screened in order to identify potentially relevant studies that satisfied our inclusion criteria. At this stage studies that clearly did not measure people’s attitudes, trust, acceptance, or anxiety toward social robots were excluded. For example, technical papers detailing the development of sensors for social robots were removed. Literature reviews, meta-analyses, editorials, newspaper articles, and other forms of popular media were also excluded at this stage as we were only interested in original empirical studies.

Second, the full-text of the identified papers were considered. Where the full-text was not available, the authors of the paper were contacted or the articles were obtained via an interlibrary loan request. Since our research questions focused on social robots exclusively, we used a pre-specified definition checklist (see Supplementary Materials 2) in order to decide whether an article was relevant or not. For example, papers investigating attitudes toward industrial robots were not included unless they also measured attitudes toward social robots. No limitations were placed on the design of the primary studies and studies with randomised and non-randomised field and lab experiments, questionnaires and surveys, interviews, pilot studies, and thesis were all included if they met the other inclusion criteria. The flow of papers through the review is detailed in Figure. 2.

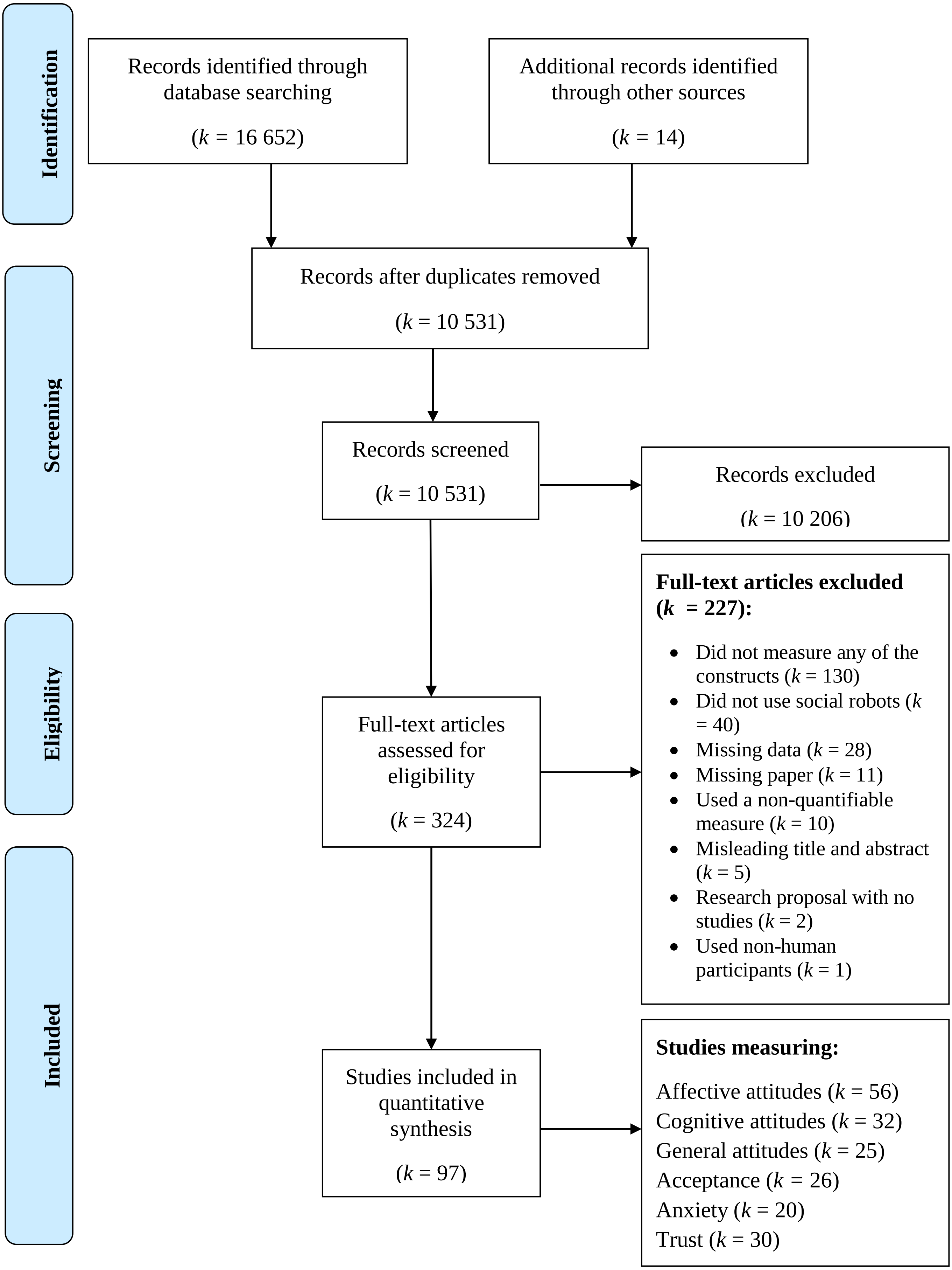


Figure 2: PRISMA flow diagram of papers through the review

3.5.3. Data extraction

The information from the primary studies was extracted by a member of the research team and 10% of the papers were second-coded by a different member of the team, with a comparison showing that 93% inter-rater agreement was reached. Any disagreements or inconsistencies between the two coders were resolved through discussion.

We first extracted bibliographic information from the articles, this included the date of publication, the country where the research was conducted, the sample size and demographics of the sample (i.e., mean age, gender, and cultural or ethnic background), the domain of application, the design of the study, and the name, design, and capabilities of the social robot. The type of outcome (categorised as general attitudes, affective attitudes, cognitive attitudes, behavioural attitudes, trust, anxiety, or acceptance) and details of the measures used to assess each outcome (e.g., the NARS) were identified and extracted next.

The methodological quality of the primary studies (i.e., risk of bias) was assessed using the tool described in Supplementary Materials 3. As with the other characteristics, a member of the research team carried out the quality assessment and a different member of the team second-coded 10% of the studies. There was moderate inter-rater agreement between the two coders, Cohen's $k = .554$, 95% CI [0.43, 0.68], $p < .001$. The average difference in the quality scores between the two coders was 0.20 points ($SD = 0.18$) for the overall methodological quality and 0.40 points ($SD = 0.16$) for the separate criteria with a maximum possible difference of 3 points. As before, disagreements were resolved via discussion.

3.5.4. Calculating and interpreting rescaled and "standardised" outcomes

Traditional effect size metrics used in meta-analyses (i.e., *r* and *d*) were not appropriate for answering our primary research question. We therefore needed a way to estimate the extent to which studies provided evidence that people have positive, neutral, or negative attitudes toward social robots. This was achieved by comparing the average value on the measure of attitude across the sample with the value of the same measure that would reflect a 'neutral' attitude (i.e., one that was neither positive nor negative). For example, if a participant completed a Likert scale measuring attitudes toward robots on a 1 to 5 scale, then a score of 3 would indicate that this participant has a neutral attitude toward robots.

In order to perform this normalisation, we calculated a pseudo-standardised sample mean ($\overline{x}_s$) and standard deviation (s_s) for each study. To calculate the pseudo-standardised scores, the mean, standard deviation, and the minimum and maximum values of each measure (i.e., scale) were identified, as well as whether the measure indicated a positive or negative outcome (e.g., whether higher values indicated a negative or a positive attitude toward robots). If a measure had multiple subscales (e.g., the NARS), then we sought to extract data separately for each subscale. Missing data was requested from authors via email or via a direct request on ResearchGate. Where the missing data was not obtained within two weeks, the papers were excluded. If articles contained multiple measures and the key statistical data was available for at least one of the measures, then the paper was included with the available data. Once all relevant data had been extracted, the following formula was used to calculate the standardised scores where $\overline{x}_s$ and s_s denote the standardised sample mean and standard deviation across participants for each study, and $\bar{x}$ and s denote the sample mean and standard deviation extracted from each study. *MR* is the mid-point of the range of a specific scale, which would denote a neutral attitude.

$$\bar{x}_s = \frac{\bar{x} - MR}{x_{max} - x_{min}} \times 2$$

$$s_s = \frac{s}{x_{max} - x_{min}} \times 2$$

Following this, an average weighted mean ($\bar{x}_w$) was calculated for each outcome. For studies that had multiple measures or subscales that assessed the same outcome (e.g., affective attitudes), the $\bar{x}_s$ and s_s for those measures were averaged. As such, each study only contributed a single $\bar{x}_s$ and s_s for a given outcome (i.e., general attitudes, affective attitudes, cognitive attitudes, behavioural attitudes, trust, anxiety, and / or acceptance). In the following formula, the mean is weighted by w_i which denotes the sample size for each study and $\sum w_i$ is the sum of all study samples for a particular outcome. We also calculated the variance ($s^2_{\bar{x}_w}$) of each weighted mean where *k* is the number of studies for each outcome, as well as the *SD* ($s_{\bar{x}_w}$), *SE* ($\sigma_{\bar{x}_w}$), and 95% Confidence Intervals where t_c is the critical *t* value for a two-tailed probability at $p < .05$.

$$\bar{x}_w = \frac{\sum_{i=1}^{n}(\bar{x}_s \times w_i)}{\sum_{i=1}^{n} w_i}$$

$$s^2_{\bar{x}_w} = \frac{\sum_{i=1}^{n} \left(w_i \times (\bar{x}_s - \bar{x}_w)^2\right)}{\frac{\sum_{i=1}^{n} w_i \times (k-1)}{k}}$$

$$s_{\bar{x}_w} = \sqrt{s^2_{\bar{x}_w}}$$

$$\sigma_{\bar{x}_w} = \frac{s_{\bar{x}_w}}{\sqrt{\frac{\sum w_i}{k}}}$$

$$95\%\ CI_{\bar{x}_w} \approx [\bar{x}_w \pm t_c \times \sigma_{\bar{x}_w}]$$

Taken together, $\bar{x}_s$ and s_s can be interpreted as a sample mean and standard deviation on a scale of -1 (indicating an extremely negative outcome) to + 1 (indicating an extremely positive outcome). Since all possible values of $\bar{x}_s$ and $\bar{x}_w$ fall within a scale with an absolute maximum and minimum values, we propose that the computed means can be interpreted in a manner that is comparable, but not identical, to that conventionally applied to Pearson's *r*. Specifically, we propose that the midpoint between neutral attitudes and the two extremes of negative and positive attitudes (i.e., $\bar{x} \geq \pm 0.50$) is interpreted as a large-sized (or *substantial*) positive or negative attitude, $\bar{x} \geq \pm 0.30$ as a medium-sized (or *moderate*) positive or negative attitude, and $\bar{x} \geq \pm 0.10$ as a small-sized (or *slight*) positive or negative attitude.

3.5.5. Calculating and interpreting weighted means, standard error, and 95% Confidence Intervals

In order to investigate whether categorical factors (e.g., type of HRI, domain of application, and robot design) are associated with people's attitudes toward social robots, we computed an average weighted mean ($\bar{x}_m$) for each level of each moderator (e.g., a weighted mean for all studies with no HRI, a weighted mean for all studies with indirect HRI, and a weighted mean for all studies with direct HRI). We excluded any studies where the outcome was measured using two or more different types of exposure to the robot, or for different robots that had different application areas, or where the outcome was measured for different robots that had different designs or no design was specified. Unlike $\bar{x}_w$, the $\bar{x}_m$ was weighted by the reported sample variance (s_s^2) in each study (in other words, we applied inverse-variance weighting instead of frequency weighting). We also calculated the variance ($s_{\bar{x}_m}^2$) of each weighted mean, as well as the *SD* ($s_{\bar{x}_m}$), *SE* ($\sigma_{\bar{x}_m}$), and 95% Confidence Intervals where t_c is the critical t value for a two-tailed probability at $p < .05$.

$$\bar{x}_m = \frac{\sum_{i=1}^{n} (x_s / s_s^2)}{\sum_{i=1}^{n} (1/s_s^2)}$$

$$s_{\bar{x}_m}^2 = \frac{1}{\sum_{i=1}^{n} (1/s_s^2)}$$

$$s_{\bar{x}_m} = \sqrt{s_{\bar{x}_m}^2}$$

$$\sigma_{\bar{x}_m} = \frac{s_{\bar{x}_m}}{\sqrt{\sum k}}$$

$$95\%\ CI_{\bar{x}_m} \approx [\bar{x}_w \pm t_c \times \sigma_{\bar{x}_m}]$$

Table 6 reports the weighed means, standard deviations, and 95% CIs for each level of each moderator. Larger positive and negative values of $\bar{x}_m$ indicate a more positive or negative outcome respectively. An overlap between confidence intervals indicates that there is insufficient evidence to conclude that there is a difference in the outcomes between the groups as a function of a given factor. Conversely, no overlap between the confidence intervals indicates that there is a difference in the outcomes between the groups as a function of a particular factor.

3.6. Results

3.6.1. Description of included studies

Data on people's acceptance of, attitudes toward, anxiety associated with, and trust in social robots was obtained from k = 97 studies published between 2005 and early 2019 in scientific journals (52%) or in conference proceedings (45%), with only three studies coming from alternative sources. The majority of these studies were conducted in the USA (17%), Germany (13%), and Japan (11%). The average size of the sample in the included studies was N = 135 (SD = 182) and the majority of studies (68%) were published between 2014 and 2019.

3.6.2. Affective attitudes

Attitudes toward social robots were most commonly assessed in terms of affective attitudes, with the majority of studies (k = 56, 58%) including at least one measure of affective attitudes (i.e., feelings or emotions toward social robots). Not surprisingly, given the popularity of the Negative Attitudes towards Robots Scale (NARS; (Nomura et al., 2006b)) in HRI research, seventeen studies (30%) used the full scale or subscales to measure participants' affective attitudes. We categorised both the NARS-S1 (interaction with robots) and NARS-S3 (emotions in interaction with robots) subscales as measures of affective attitudes, as the items enquire how people expect to feel when they interact with social robots. Other measures of affective attitudes included other validated scales (e.g., Godspeed Questionnaire Series – likability) and less-known self-report measures (e.g., semantic differential scales based on Crites et al., 1994). Twelve studies (21%) measured participants' affective attitudes towards robots in general (e.g., Dinet, & Vivian, 2014) or types of robots (e.g., domestic robots; (de Graaf et al., 2016)), while the rest measured participants' attitudes towards individual robots (e.g., NAO; (Torta et al., 2014)).

The average weighted mean for affective attitudes was $\bar{x}_w$ = 0.27 (see Figure 3), suggesting that people generally have slight (bordering on moderate) positive affective attitudes toward social robots. Eight studies (14%) found evidence that people held negative affective attitudes toward social robots (i.e., $\bar{x}_w < 0$) and only 16 studies (29%) had a mean of $\bar{x}_s > \pm 0.50$, signifying that people held substantially positive or negative affective attitudes.

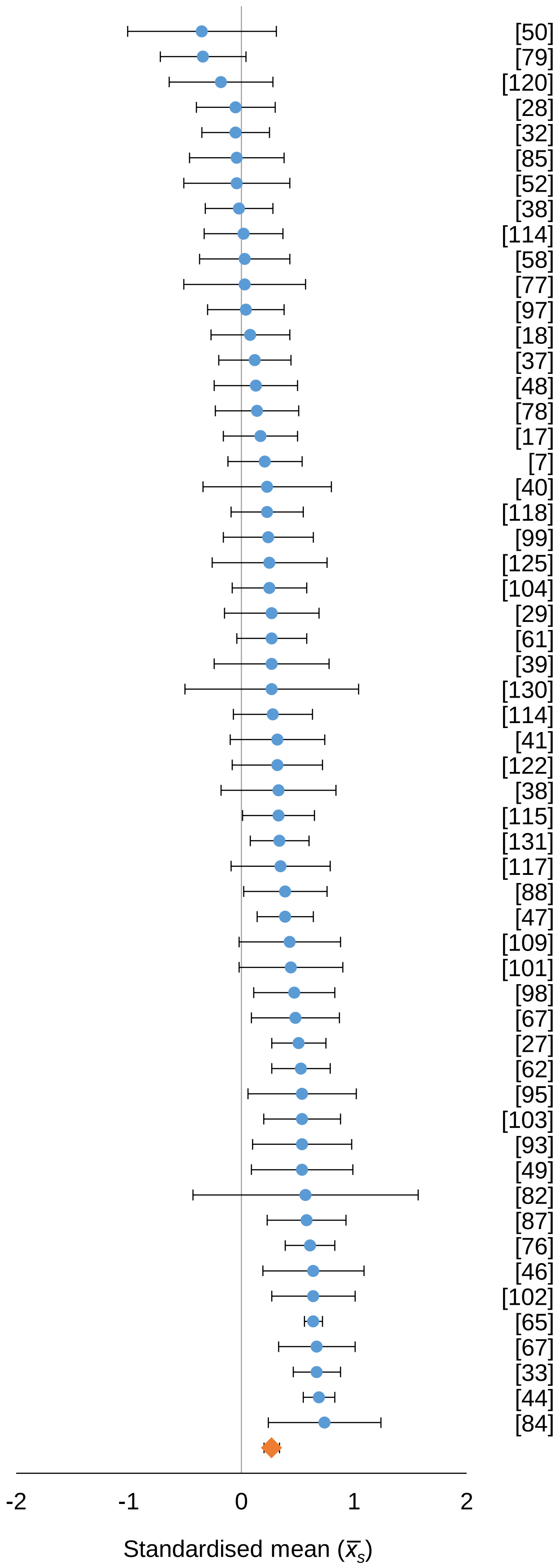


Figure 3: Plot of pseudo-standardised means ($\overline{x}_s$) for studies measuring affective attitudes toward social robots. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\overline{x}_w$) for affective attitudes and the error bars represent **95%** $\boldsymbol{CI}_{\overline{x}_w}$

3.6.3. Cognitive attitudes

Thirty-two studies (33%) included at least one measure of cognitive attitudes (i.e., people's cognitive evaluations or thoughts about social robots). The NARS, or more specifically the NARS-S2 subscale (reflecting beliefs about the social influence of robots), was the most commonly used measure (k = 17, 53%). Subscales of questionnaires relating to specific models such as the Almere Model of robot acceptance (Heerink et al., 2010) and Unified Theory of Acceptance and Use of Technology (UTAUT; (Venkatesh et al., 2003)) were also used to measure cognitive attitudes (Conti, Di Nuovo, Buono, & Di Nuovo, 2017; Shin, & Choo, 2011; Tay, Jung, & Park, 2014).

The average weighted mean for cognitive attitudes was $\bar{x}_w$ = 0.24, indicating that, in general, people had slightly positive beliefs about robots and their use (see Figure 4). The majority of studies (72%) found evidence for positive cognitive attitudes with one study, (Nomura, 2014), providing evidence for neutral cognitive attitudes ($\bar{x}_s \approx 0$).

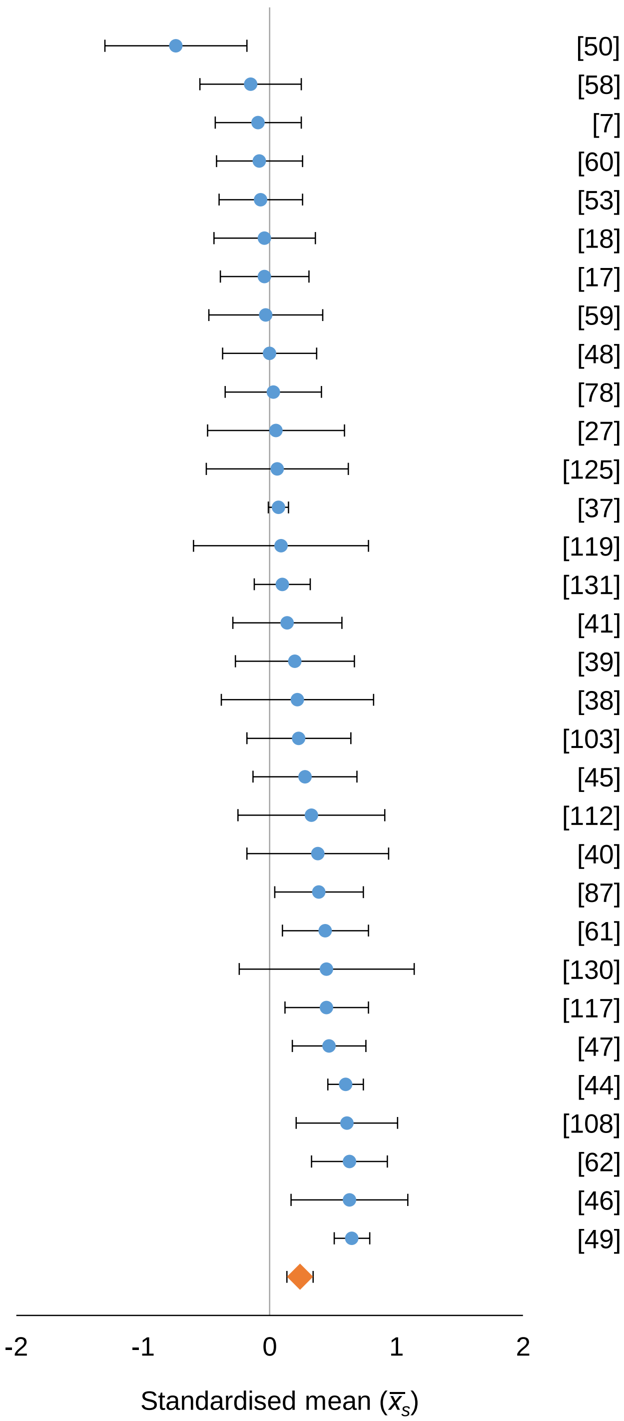


Figure 4: Plot of pseudo-standardised means ($\bar{x}_s$) for studies measuring cognitive attitudes toward social robots. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\bar{x}_w$) for cognitive attitudes and the error bars represent $\mathbf{95\%}\ \boldsymbol{CI}_{\bar{x}_w}$

3.6.4. General attitudes

Twenty-five studies (26%) measured attitudes towards social robots in a general way – i.e., overall evaluations of the extent to which social robots are ‘good’ or ‘bad’ and / or measures that combined affective and cognitive evaluations. General attitudes were almost exclusively measured via self-report with the exception of studies using the Implicit Association Test (IAT). The aggregated data (see Figure 5) indicated an average weighted mean of $\bar{x}_w = 0.07$, which suggests that people’s general attitudes towards social robots tended to be neutral (bordering on slightly positive). Thirteen studies (55%) provided evidence of positive general attitudes (i.e., $\bar{x}_w > 0$) towards social robots while the rest provided evidence for negative attitudes, with one study, reporting neutral attitudes (i.e., $\bar{x}_w = 0$).

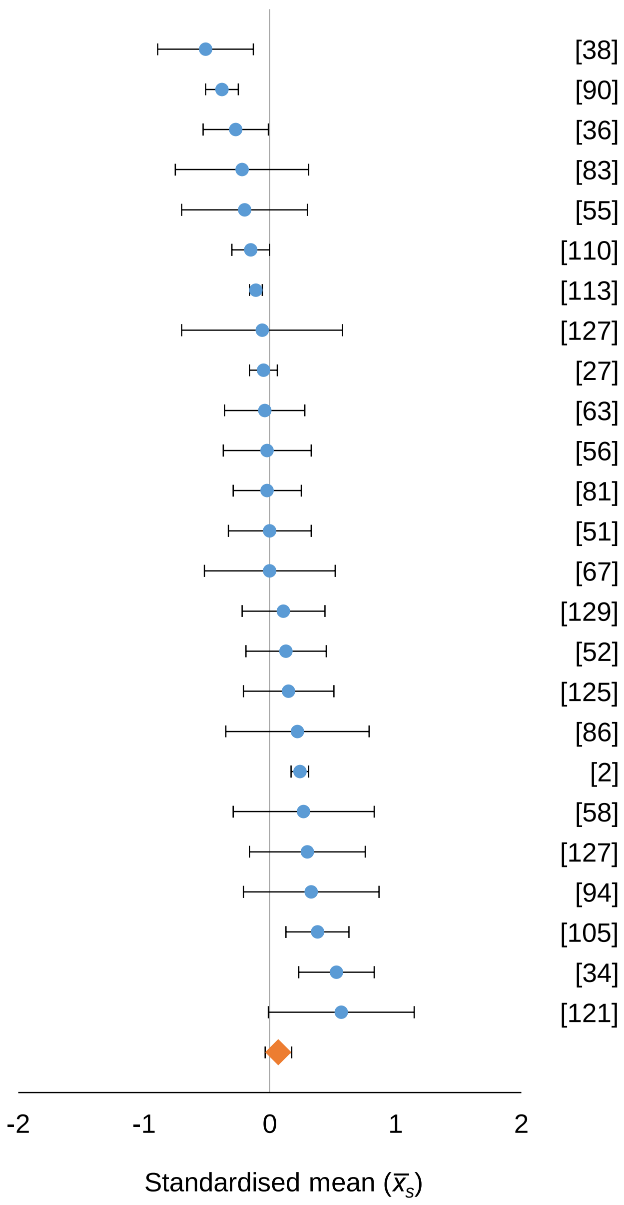


Figure 5: Plot of pseudo-standardised means ($\bar{x}_s$) for studies measuring general attitudes toward social robots. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\bar{x}_w$) for general attitudes and the error bars represent $\mathbf{95\%}\ \boldsymbol{CI}_{\bar{x}_w}$

3.6.5. Acceptance

Twenty-six of the included studies (27%) measured acceptance in terms of people's intentions to use social robots, actual use of specific social robots or social robots in general, or people's willingness to interact with social robots. The average weighted mean for this outcome ($\bar{x}_w = 0.24$) indicated that, in general, people accept social robots but only slightly so. However, acceptance of social robots varied considerably (see Figure 6) and 42% of studies suggested that people did not accept robots (i.e., $\bar{x}_w < 0$).

Two studies in particular should be mentioned as they are rather atypical as compared to the other studies measuring acceptance. First, Fridin and Belokopytov (2014) reported an unusually small standard deviation (s_s) indicating very little variation in participants' acceptance of social robots. This may be explained by the specific conditions and sample in this study. Participants were all preschool and elementary school teachers that attended a professional workshop on educational robotics where they were introduced to the capabilities of a NAO robot. This may explain why participants' views on robots aligned quite well. Second, Wu et al. (2014) found strong evidence that participants did not accept robots ($\bar{x}_w = -0.99$) These negative beliefs may be explained by the finding that the participants who interacted with a social robot for a month in a Living Lab setting did not find the robot useful. Perceived usefulness has previously been identified as a factor that impacts participants' intention to use robots (Venkatesh et al., 2003).

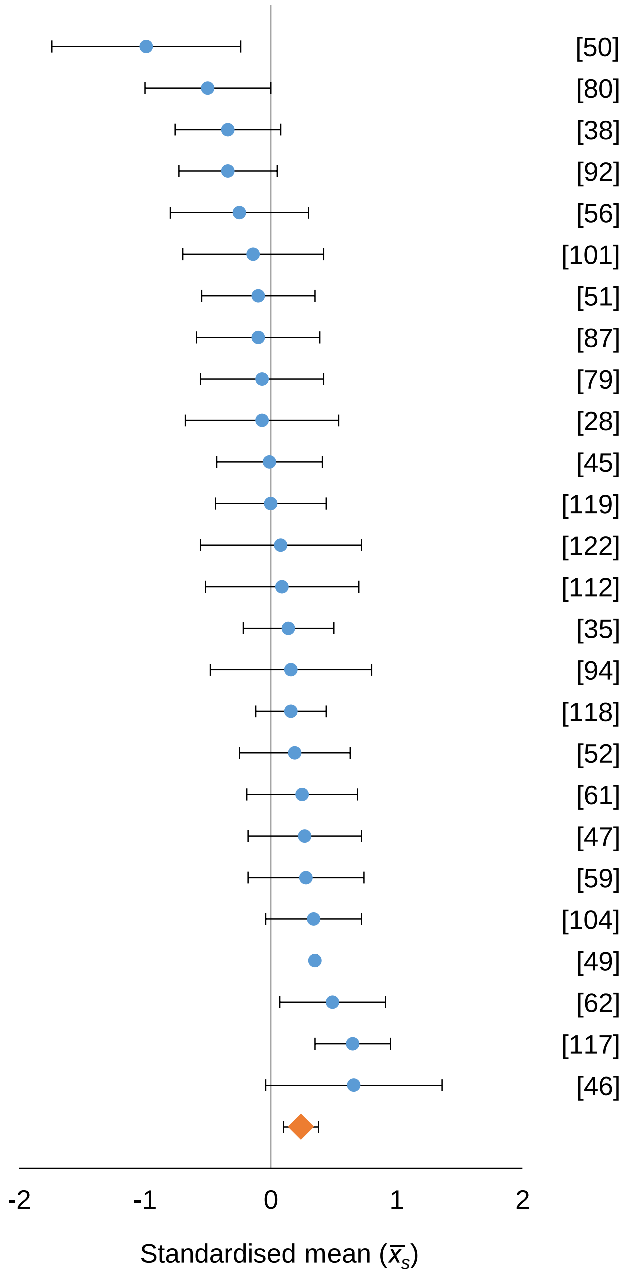


Figure 6: Plot of pseudo-standardised means ($\bar{x}_s$) for studies measuring acceptance toward social robots. Positive values represent greater acceptance. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\bar{x}_w$) for acceptance and the error bars represent $\mathbf{95\%}\ \boldsymbol{CI}_{\bar{x}_w}$

3.6.6. Anxiety

Twenty studies (21%) measured people's feelings of anxiety or nervousness evoked by social robots. Anxiety was predominantly assessed via the Robot Anxiety Scale (RAS; (Nomura et al., 2006c)) with ten studies (50%) having used some variation of the measure (de Graaf et al., 2016; Kuchenbrandt, & Eyssel, 2012; Wullenkord, & Eyssel, 2014). Other commonly used measures (k = 5, 25%) were the subscales of adapted questionnaires relating to specific models such as the Almere Model of robot acceptance (Heerink et al., 2010) and Unified Theory of Acceptance and Use of Technology (UTAUT; (Venkatesh et al. 2003)). All of the studies used self-report measures of anxiety with some studies measuring either anxiety toward specific social robots or toward social robots in general.

We found an average weighted mean of $\bar{x}_w = 0.10$ for anxiety, indicating that, in general, people only feel slightly anxious about social robots. Indeed, the majority of studies ($k = 9$, 45%) found that participants' levels of anxiety were fairly neutral (i.e., $\bar{x}_w < \pm0.10$, see Figure 7 6). The 95% $CI_{\bar{x}_w}$ further support this conclusion with confidence limits that cross 0 but do not exceed $\bar{x}_w = -0.10$ (see Figure 7).

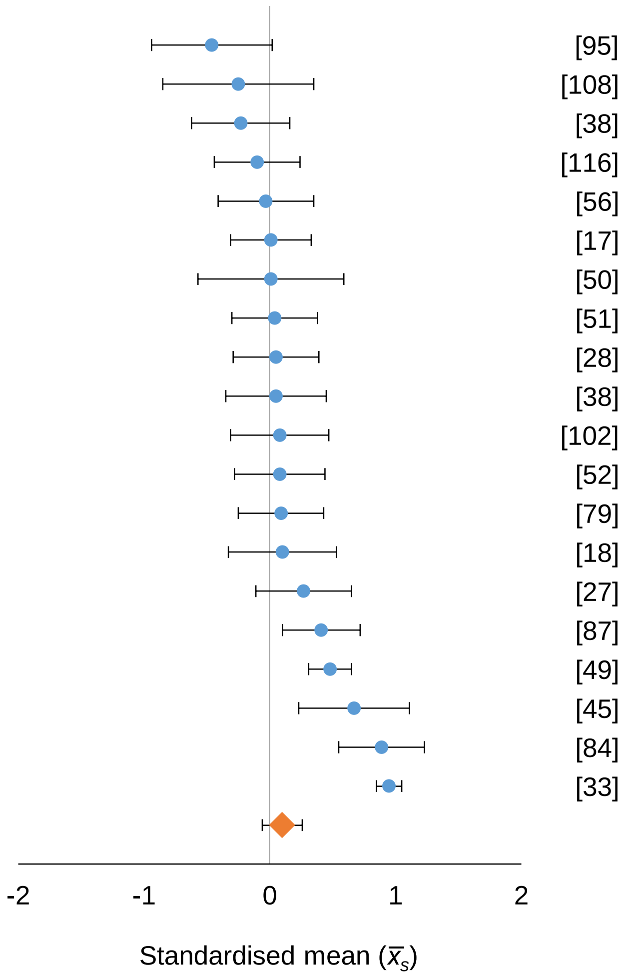


Figure 7: Plot of pseudo-standardised means ($\bar{x}_s$) for studies measuring anxiety toward social robots. Positive values represent lesser anxiety. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\bar{x}_w$) for anxiety and the error bars represent $\mathbf{95\%}\ \boldsymbol{CI}_{\bar{x}_w}$

3.6.7. Trust

Thirty studies (31%) measured trust in social robots. Unlike the other outcomes, measures of trust were notably more varied and included behavioural (Gaudiello et al., 2016; Stanton, & Stevens, 2017) as well as self-report measures. However, trust was typically assessed via subscales of adapted questionnaires relating to specific models such as the Almere Model of robot acceptance (Heerink et al., 2010) and the Unified Theory of Acceptance and Use of Technology (Venkatesh et al. 2003).

The average weighted mean for trust was close to zero, $\bar{x}_w$ = 0.06, suggesting that, in general, people did not particularly trust or distrust social robots. However, the plot of all included studies (see Figure 8) indicated variation within and between studies with 43% of studies presenting evidence that people did not trust social robots (i.e., $\bar{x}_w < 0$).

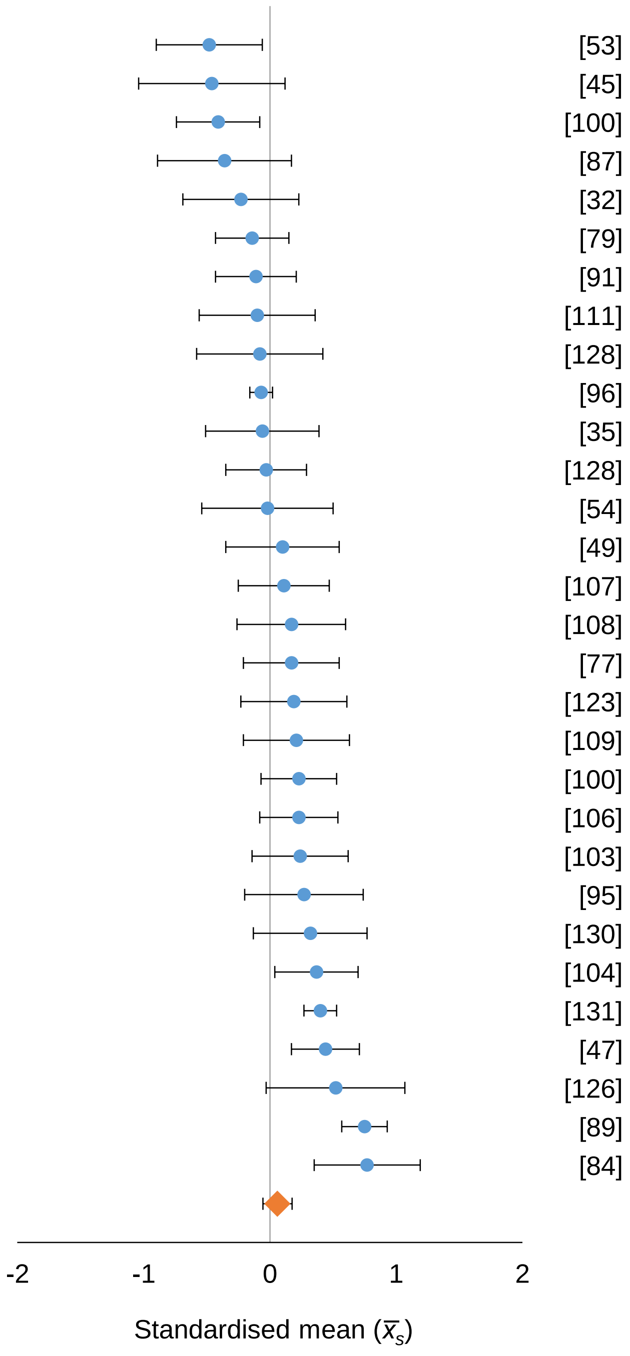


Figure 8: Plot of pseudo-standardised means ($\bar{x}_s$) for studies measuring trust toward social robots. Positive values represent greater trust. Error bars of the blue data points represent the standard deviation (s_s) of the mean. The orange data point represents the average weighted mean ($\bar{x}_w$) for trust and the error bars represent $\mathbf{95\%}\ \boldsymbol{CI}_{\bar{x}_w}$

3.6.8. Factors that influence the main outcomes

Table 6 shows the weighted means ($\bar{x}_m$) and confidence intervals ($95\%\ CI_{\bar{x}_m}$) for each outcome as a function of factors that might influence that outcome (e.g., the nature of the social robot). In addition, the findings have been illustrated graphically in Supplementary Materials 1.

3.6.8.1. Type of exposure to robots

We compared attitudes in studies that included three different types of human-robot interaction: no HRI, an indirect form of HRI, and direct HRI. For studies measuring affective attitudes, the average weighted mean for studies that did not include any type of HRI was larger ($\bar{x}_m$ = 0.40) than for studies where indirect contact ($\bar{x}_m$ = 0.09) with social robots was included. We also found more positive affective attitudes toward social robots for studies that included direct HRI ($\bar{x}_m$ = 0.34) as compared to indirect HRI ($\bar{x}_m$ = 0.09). There was no evidence that affective attitudes differed between no HRI and direct HRI ($\bar{x}_m$ = 0.34). This suggests that, in general, when people are asked about their feelings toward social robots, they report more positive affective attitudes when they either do not interact with a social robot at all or directly interact with it, rather than when they experience some type of indirect contact.

For cognitive attitudes, there was no overlap between the $95\%\ CI_{\bar{x}_m}$ for no HRI and direct HRI, indicating that participants thoughts about social robots were more positive in studies where there was no interaction between participants and robots ($\bar{x}_m$ = 0.35) than when there was direct interaction ($\bar{x}_m$ = -0.13). There was no evidence that cognitive attitudes differed between studies that involved indirect HRI ($\bar{x}_m$ = 0.37) and no HRI.

With respect to general attitudes, participants appeared to report more positive attitudes toward social robots in studies with indirect forms of HRI ($\overline{x}_m$ = 0.22) than in studies with direct ($\overline{x}_m$ = -0.14) or no HRI ($\overline{x}_m$ = -0.10). This lack of overlap between the 95% $CI_{\overline{x}_m}$ suggests that participants attitudes toward social robots tend to be more positive when they interact with the robots indirectly (e.g., by watching a video; (Cramer et al., 2009)) rather than when they interact directly or do not interact with a social robot at all.

There was no overlap between confidence intervals for acceptance of social robots between studies where there was no HRI ($\overline{x}_m$ = 0.42) and for studies with indirect HRI ($\overline{x}_m$ = -0.14), suggesting that, in general, people are more accepting of social robots with which they have had no contact as compared to robots they have interacted with indirectly.

For anxiety, there was no overlap between confidence intervals for studies that included direct and indirect HRI. This indicates that, in general, participants reported considerably less anxiety when directly interacting with social robots ($\overline{x}_m$ = 0.65) than when taking part in indirect HRI ($\overline{x}_m$ = 0.03) or no HRI ($\overline{x}_m$ = 0.10).

Results from the studies measuring trust were consistent with the findings for anxiety. In general, for studies where there was direct HRI, participants exhibited or reported more trust in social robots ($\overline{x}_m$ = 0.18) than participants in studies where the contact with the social robots was indirect ($\overline{x}_m$ = -0.06). Unfortunately, too few studies measured trust in the absence of HRI so we were unable to compare this group to indirect and direct HRI.

In addition to considering whether the type of exposure to robots provided in experimental studies influences people's beliefs about robots, we also sought to examine the effects of long-term exposure to robots by comparing attitudes and beliefs in studies where the majority (i.e., over half) of the participants indicated that they had seen or interacted with robots with studies where more than half of the participants had not previously seen or interacted with robots. Although fourteen studies reported the number of participants that had seen or interacted with social robots previously, in all but one of those studies the majority of participants had no previous experience with robots. Therefore, it was not possible to examine the effect of long-term interactions on beliefs about social robots in this review.

3.6.8.2. Domain of application

We looked at attitudes toward robots in six different domains of application: (i) companionship and domestic use, (ii) education, (iii) general application, (iv) healthcare, (v) HRI, and (vi) paediatric care.

We found three main differences, indicated by no overlap between confidence intervals, for studies measuring affective attitudes. In general, participants' affective attitudes toward social robots intended for companionship or domestic purposes were more positive ($\bar{x}_m$ = 0.45) than were participants' attitudes toward social robots intended to have a general application ($\bar{x}_m$ = 0.13). In addition, participants had more positive affective attitudes toward social robots in healthcare settings ($\bar{x}_m$ = 0.58) than robots with a general or HRI-focused application ($\bar{x}_m$ = 0.13 and 0.34, respectively).

Participants reported more positive cognitive attitudes toward social robots in educational domains ($\bar{x}_m$ = 0.59) than did participants where the social robot had a general ($\bar{x}_m$ = 0.07) or HRI-focused ($\bar{x}_m$ = 0.12) application. There were no other differences of note.

No differences in general attitudes were found as a function of the domain of application as the confidence intervals for all groups overlapped. However, it should be noted that we could only identify enough studies to compare general attitudes toward social robots in three domains of application – healthcare, general application, and HRI.

With respect to acceptance, participants seemed more accepting of social robots in educational domains ($\bar{x}_m$ = 0.35) than social robots with a general, healthcare, or HRI-focused application ($\bar{x}_m$ = 0.07, 0.02, and -0.02, respectively).

We were only able to compare three different domains of application for studies measuring anxiety and found no evidence of differences in anxiety associated with social robots as a function of their domain of application.

Finally, we compared trust associated with social robots in three domains of application. There was a difference in trust between studies where the social robot had a healthcare application and studies where the social robot had an HRI-focused application as indicated by no overlap between the confidence intervals for those two groups. Participants reported less trust in social robots intended for healthcare settings ($\bar{x}_m$ = 0.09) and for general application ($\bar{x}_m$ = -0.04), than in social robots intended for HRI ($\bar{x}_m$ = 0.32).

3.6.8.3. Design of robot

We looked at differences between three broad categories of social robots' design: anthropomorphic, humanoid, and non-humanoid robots. Unfortunately, for all six outcomes, the majority of studies focused exclusively on participants' attitudes toward humanoid social robots (see Table 6). As such, there was insufficient evidence on people's beliefs about anthropomorphic and non-humanoid social robots, resulting in fairly large confidence intervals that made comparisons difficult. Consequently, we were either unable to compare the three design groups or found no evidence of differences in affective attitudes, cognitive attitudes, acceptance, anxiety, general attitudes, or trust as a function of the design of the social robot.

3.6.8.4. Geographical location

We sought to compare attitudes between eight geographical locations in which the data collection took place: Australia, France, Germany, Italy, Japan, Netherlands, New Zealand, South Korea, Taiwan, and the USA (see Table 6).

We found three main differences, indicated by no overlap between confidence intervals, for studies measuring affective attitudes. In general, estimates of participants' affective attitudes toward social robots from studies conducted in Italy ($\bar{x}_m$ = 0.57) were more positive than were participants' attitudes from studies conducted in Germany ($\bar{x}_m$ = 0.22), Japan ($\bar{x}_m$ = 0.21), and the USA ($\bar{x}_m$ = 0.05).

Participants from studies conducted in France ($\bar{x}_m$ = 0.35) reported more positive cognitive beliefs about social robots than did participants who took part in studies conducted in Japan ($\bar{x}_m$ = 0.05). No other differences between people's cognitive attitudes were found, although it should be noted that due to a limited number of studies we were only able to compare four of the eight eligible geographical locations.

We were only able to compare people's general attitudes from studies conducted in Germany, the Netherlands, New Zealand, and the USA. We found that participants' general attitudes toward social robots from studies conducted in New Zealand ($\bar{x}_m$ = 0.23) tended to be more positive than those of studies conducted in the USA ($\bar{x}_m$ = -0.10).

We were unable to compare acceptance between the countries as Germany was the only geographical location for which we had enough data to calculate $\bar{x}_m$.

We found no differences in the levels of anxiety people experience toward social robots as a function of the location at which the study was conducted. However, we were only able to compare studies conducted in Germany, Italy, Japan, and the Netherlands, and on average we were only able to include three studies per location resulting in large 95% CIs that made comparisons difficult.

Similarly, we found no differences in people's level of trust in social robots as a function of the location at which the study was conducted and we were only able to compare studies conducted in Australia, Italy, and the USA.

3.6.8.5. Age of participants

In order to investigate whether participants' age was associated with their beliefs about social robots, we conducted a weighted least squares regression with the average age of participants in each study as the independent variable, the sample mean ($\bar{x}_s$) as the dependant variable, and the size of the sample in each study as the weight. A Bonferroni correction was applied to account for the multiple comparisons and an adjusted critical p value of .008 was used. These analyses indicated that the age of the participants was not significantly associated with their affective attitudes toward social robots, $F(1, 43) = 1.90$, $p = .176$, cognitive attitudes toward social robots, $F(1, 22) = 1.90$, $p = .182$, general attitudes, $F(1, 22) = 0.00$, $p = .948$, acceptance, $F(1, 21) = 3.80$, $p = .065$, anxiety, $F(1, 16) = 0.00$, $p = .981$, or trust, $F(1, 20) = 1.35$, $p = .259$.

3.6.8.6. Gender of participants

In order to investigate whether gender was associated with participants' beliefs about social robots, we conducted a weighted least squares regression with the percentage of female participants in each study as the independent variable, the sample mean ($\overline{x}_s$) as the dependant variable, and the size of the sample in each study as the weights. A Bonferroni correction was applied to account for the multiple comparisons and an adjusted critical p value of .008 was used. The percentage of female participants accounted for 40.9% of the variation in self-reported trust in social robots, $R^2 = .64$, $F(1, 19) = 13.16$, $p = .002$, such that there was a strong positive linear relationship between the two. However, the gender of the participants was not associated with their affective attitudes toward social robots, $F(1, 45) = 1.98$, $p = .166$, cognitive attitudes, $F(1, 24) = 0.04$, $p = .853$, general attitudes, $F(1, 20) = 4.28$, $p = .052$, acceptance, $F(1, 20) = 5.70$, $p = .658$, or anxiety, $F(1, 13) = 5.89$, $p = .031$.

3.6.8.7. Year of publication

In order to investigate whether beliefs about social robots have changed over time, we conducted a weighted least squares regression for each of the six outcomes with the year in which the study was published as the independent variable, the sample mean ($\overline{x}_s$) as the dependant variable, and the sample size of each study as the weight. The average number of studies published each year prior to 2014 was quite small ($M = 3.44$) and therefore the findings of the linear regressions should be interpreted with caution. The year of publication was not associated with affective attitudes, $F(1, 55) = 0.17$, $p = .684$; cognitive attitudes, $F(1, 31) = 0.49$, $p = .489$; general attitudes, $F(1, 23) = 3.00$, $p = .096$; acceptance, $F(1, 23) = 0.32$, $p = .575$; anxiety, $F(1, 18) = 0.03$, $p = .856$; or trust, $F(1, 28) = 0.001$, $p = .986$.

3.6.8.8. Methodological quality

The average overall methodological quality of the included studies was 2.20 (*SD* = 0.50, range = 1.30–3.30) on a scale from 1 (poor quality) to 4 (excellent quality) (see Supplementary Materials 3). It should be noted that most studies received a quality score close to the average, indicating little variation in the overall methodological quality as measured via our Quality Assessment Tool. However, a number of individual criterion may have contributed to this homogeneity. Most notably, the Objectivity criterion (*M* = 2.00, *SD* = 0.20) as the majority of studies (94%) measured our main outcomes using some form of questionnaire or scale which we rated as lower than behavioural and physiological measures. Similarly, the Reliability (a) criterion (*M* = 1.30, *SD* = 0.60) indicated that the majority of studies did not measure test-retest reliability, thus resulting in a score of 1 for the majority of studies (70%). Scores for the External Validity (b) criterion were similarly homogeneous (*M* = 1.40, *SD* = 0.60) as most studies did not employ a randomised sampling technique. By far the most common type of sample used by 30% of the studies consisted of University students recruited on a volunteer basis.

In order to investigate whether the methodological quality of studies was associated with participants' beliefs about social robots, we conducted a Linear Regression with the methodological quality scores of each study as the independent variable, and the sample mean ($\overline{x}_s$) as the dependant variable. A Bonferroni correction was applied to account for the multiple comparisons and an adjusted critical p value of .008 was used. The methodological score given to the included studies was not associated with participants' affective attitudes toward robots, $F(1, 54) = 1.25$, $p = .269$; cognitive attitudes, $F(1, 30) = 0.02$, $p = .878$; general attitudes, $F(1, 23) = 2.39$, $p = .136$; acceptance of robots, $F(1, 24) = 1.33$, $p = .260$; anxiety toward robots, $F(1, 18) = 1.19$, $p = .056$; and trust in robots, $F(1, 31) = 0.37$, $p = .549$.

Table 6

Weighted means ($\bar{x}_m$), weighted standard deviations ($s_{\bar{x}_m}$), total sample size (N), number of studies (k), and weighted 95% Confidence Intervals for outcomes as a function of factors that might influence outcomes.

	Affective attitudes					Cognitive attitudes					General attitudes				
	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]
Type of HRI															
No HRI	0.46	0.09	4348	10	[0.40, 0.52]	0.40	0.10	4506	9	[0.33, 0.48]	-0.10	0.04	1854	6	[-0.14, -0.05]
Indirect HRI	0.09	0.16	1063	7	[-0.05, 0.24]	0.37	0.26	558	4	[-0.05, 0.78]	0.22	0.06	1544	8	[0.16, 0.27]
Direct HRI	0.33	0.05	1982	29	[0.31, 0.34]	-0.11	0.08	1196	13	[-0.15, -0.06]	-0.14	0.09	838	8	[-0.22, -0.07]
Area of robot application															
Companionship and Domestic	0.45	0.13	703	4	[0.23, 0.66]	-	-	674	2	-	-	-	384	2	-
Education	0.23	0.21	832	4	[-0.11, 0.57]	0.59	0.12	652	3	[0.28, 0.90]	-	-	375	2	-
General application	0.13	0.07	4171	15	[0.09, 0.17]	0.07	0.08	4160	10	[0.02, 0.13]	-0.09	0.11	2389	7	[-0.20, 0.01]
Healthcare	0.58	0.07	563	7	[0.51, 0.65]	0.09	0.25	282	4	[-0.30, 0.48]	-0.02	0.04	660	6	[-0.06, 0.02]
HRI	0.34	0.07	1351	22	[0.31, 0.37]	0.12	0.11	845	10	[0.04, 0.20]	-0.06	0.14	562	7	[-0.19, 0.06]
Paediatric care	0.36	0.30	235	3	[-0.37, 1.10]	-	-	188	2	-	-	-	172	1	-
Design of robot															
Anthropomorphic	0.24	0.17	286	6	[0.07, 0.42]	0.24	0.24	141	3	[-0.36, 0.84]	-	-	57	1	-
Humanoid	0.34	0.05	1253	26	[0.32, 0.35]	0.10	0.07	1040	15	[0.06, 0.13]	-0.18	0.09	625	9	[-0.25, -0.11]
Non-humanoid	0.33	0.17	856	5	[0.11, 0.55]	-	-	467	1	-	-	-	41	1	-

Table 6 (continued)

	Acceptance					Anxiety					Trust				
	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]	$\bar{x}_m$	$s_{\bar{x}_m}$	N	k	95% $CI_{\bar{x}_m}$ [LL, UL]
Type of HRI															
No HRI	0.42	0.20	2168	4	[0.10, 0.74]	0.11	0.23	873	3	[-0.45, 0.68]	-	-	24	1	-
Indirect HRI	-0.14	0.20	608	6	[-0.35, 0.08]	0.03	0.19	394	4	[-0.27, 0.33]	-0.06	0.15	574	6	[-0.22, 0.09]
Direct HRI	0.03	0.15	671	10	[-0.08, 0.14]	0.63	0.08	374	10	[0.57, 0.69]	0.18	0.06	895	20	[0.16, 0.21]
Area of robot application															
Companionship and Domestic	-	-	-	1	-	-	-	215	2	-	-	-	29	2	-
Education	0.35	0.02	1152	4	[0.32, 0.38]	0.34	0.14	543	3	[-0.02, 0.69]	-	-	18	1	-
General application	0.07	0.19	1526	6	[-0.13, 0.27]	-	-	2	1	-	-0.04	0.08	282	6	[-0.13, 0.04]
Healthcare	0.02	0.28	260	4	[-0.42, 0.46]	0.36	0.22	100	4	[0.01, 0.71]	0.09	0.16	256	5	[-0.11, 0.29]
HRI	-0.02	0.15	913	8	[-0.15, 0.10]	0.05	0.13	760	8	[-0.05, 0.16]	0.32	0.08	884	13	[0.27, 0.37]
Paediatric care	-	-	88	2	-	-	-	80	1	-	-	-	88	2	-
Design of robot															
Anthropomorphic	-0.08	0.30	195	3	[-0.83, 0.67]	0.26	0.19	153	4	[-0.03, 0.56]	0.25	0.14	327	6	[0.10, 0.40]
Humanoid	0.34	0.02	1037	12	[0.33, 0.35]	0.62	0.07	571	10	[0.57, 0.67]	0.14	0.06	919	18	[0.11, 0.17]
Non-humanoid	-	-	182	2	-	-	-	83	2	-	-	-	150	2	-

3.7. Discussion

The present review quantified and synthesised evidence on people's attitudes towards, anxiety associated with, trust in, and acceptance of social robots. Although reviews have been conducted in this area (Hancock et al., 2011a; Broadbent et al., 2009; Savela et al., 2018; Chen, & Chan, 2011), none have combined the various measures employed in primary studies in a way that informs the overall valence (i.e., positive, neutral, or negative) and magnitude of the outcomes. The approach described in this paper is, to our knowledge, the first of its kind to provide standardised estimates of the overall valence of people's attitudes toward robots and related beliefs based on evidence from multiple studies and measures.

3.7.1. What are people's attitudes towards social robots?

The majority of studies that measured people's affective attitudes suggested that people have slightly positive (bordering on moderate) feelings about social robots. We consider this finding to be fairly robust as only nine studies provided evidence that people have negative feelings toward social robots. Upon further examination of these nine studies, two had somewhat atypical methodologies – one study employed imagined contact with robots (Kuchenbrandt, & Eyssel, 2012) and the other tested whether involving users in the development of robots affected their attitudes (Reich-Stiebert, & Eyssel, 2015).

Studies measuring cognitive attitudes provided further support for overall positive attitudes toward robots with a sample-weighted mean similar in magnitude to that found for affective attitudes. This similarity between affective and cognitive attitudes is consistent with models in psychology that propose a moderate correlation between the three components of attitude (Breckler, 1984; Ostrom, 1969). However, it is possible for there to be differences between what people feel and think about specific robots, as is the case for some of the studies included in the present review (Nomura et al., 2006a; Backonja et al., 2018; Rantanen, Lehto, Vuorinen, & Coco, 2018). The impact of dissonance between affective and cognitive attitudes in relation to human-robot interaction has not yet been investigated and warrants consideration.

Where studies used measures of attitude that did not reflect purely affective or cognitive attitudes, or it was not possible to obtain data for subscales measuring different outcomes (e.g., the NARS), we coded said measures under the blanket term of general attitudes. Findings for this outcome were not entirely consistent with the results for affective and cognitive attitudes, as the sample-weighted mean was almost zero and thus indicated a relatively neutral rather than slightly positive attitude. Indeed, compared to the other outcomes, the number of studies providing evidence for negative attitudes was much greater (i.e., approximately half of the studies). It is possible that this finding was a product of some difference in the methodology or measures that necessitated the studies' inclusion in the general category. For example, NARS subscales may have been combined if the reliability of the subscales was poor.

Although we coded the outcomes in the primary studies based on definitions rooted in social psychological research on attitudes (see section 3.2), it should be noted that studies generally did not differentiate between the various types of attitudes and often did not provide a definition of attitudes at all. This may be of some concern especially if it indicates a poor understanding of the relationship between attitudes and behaviour. Given the number of studies that measured attitudes in the context of human-robot interaction (Zlotowski et al., 2015; de Graaf, & Allouch, 2013, Wullenkord, Fraune, Eyssel, & Šabanović, 2016) and sometimes with the purpose of predicting behaviour (de Graaf, & Allouch, 2013; Nomura et al., 2008; Park, &

Del Pobil, 2012; Spence, Edwards, & Edwards, 2018; Wullenkord et al., 2016; Zlotowski et al., 2015), attitude-behaviour models from social psychology should be used more consistently to inform HRI research (Hewstone, & Swart, 2011; Pettigrew, 1998; Pratto, Sidanius, & Levin, 2006).

3.7.2. To what extent do people accept, trust, and feel anxious towards robots?

We found that, in general, people are either willing to use social robots or have the intention to do so given the chance. Given the conceptual overlap between acceptance of social robots and behavioural attitudes, it is not surprising that our findings with respect to acceptance are similar to our findings for affective and cognitive attitudes. This is once again consistent with research supporting a moderate correlation between the three components of attitude (Breckler, 1984; Ostrom, 1969).

Findings from the studies measuring trust indicated that, in general, people neither explicitly trusted or mistrusted robots; rather they typically were neutral with respect to trust. However, given the variability in estimates of trust across studies (i.e., some studies reported high trust and others low trust) it is likely that the extent to which people trust social robots is moderated by other factors, some of which we discuss below in section 3.7.3.

Finally, we found evidence suggesting that people are fairly neutral in terms of the anxiety that they report with respect to social robots. This finding may, to a certain extent, be a product of the general tendency for social robots to be designed in such a way as to appear less threatening. For example, NAO, a generally well-liked robot (Hosseini et al., 2017; Von Der Pütten, & Krämer, 2014; Torta, Oberzaucher, Werner, Cuijpers, & Juola, 2013), was used in 45% of the studies measuring anxiety and may have contributed to the overall neutral to positive valence for anxiety and trust.

3.7.3. What factors affect the main outcomes?

We found mixed evidence that exposure to robots, domain of application and design of the robots, and the age and gender of participants was associated with people's beliefs about robots. This was predominantly due to a limited number of studies which meant that it was not possible to reliably estimate beliefs for the different categories of many of the factors of interest. Indeed, affective attitudes was the only outcome for which it was possible to compare all categories across all the factors. Additionally, whether participants were exposed to robots (directly or indirectly) before their beliefs were measured was the only factor for which comparison across the outcome measures was possible. As such we will focus on these findings first.

We found mixed evidence on whether and how exposure to robots affects people's attitudes and beliefs. Participants typically reported positive affective attitudes regardless of whether they interacted with a robot or not. However, people's affective attitudes toward social robots in studies with indirect HRI were typically less positive than participants' affective attitudes in studies with no HRI or direct HRI. This suggests that interacting with a robot face-to-face elicits more positive feelings toward said robot (or robots in general) than does some form of indirect contact such as watching a video of the robot. These findings may be an important consideration when measuring attitudes in HRI contexts where the affective evaluation of a robotic platform during indirect contact may not accurately represent people's feelings toward that robot, or social robots in general (Wullenkord et al., 2016; Bazzano, & Lamberti, 2018).

Notably, interaction did not seem to have the same effect on cognitive or general attitudes. For example, studies involving direct contact typically found that people held negative cognitive and general attitudes toward social robots. This finding is somewhat contrary to assertions that directly interacting with robots is a potential strategy for improving attitudes toward them (Bartneck et al., 2007a; Wullenkord et al., 2016, Stafford et al., 2010). It could be that while the novelty of directly interacting with a social robot results in positive affect it also allows participants to identify potential issues with robotic platforms or make general observations about their usefulness that result in negative thoughts. Supporting this idea is our finding that, unlike affective attitudes, participants typically reported more positive cognitive and general attitudes in studies utilising indirect contact (where it could be more difficult to identify issues with robotic platforms) than in studies with direct HRI. Due to a lack of studies utilising contact other than direct HRI, it was not possible to draw definitive conclusions regarding the impact of exposure to robots on people's acceptance of, anxiety toward, and trust in social robots.

Although we found some differences in participants' affective, cognitive, and general attitudes between geographical locations, these findings were limited by the number of studies available for comparison for nearly all outcomes. This was partly due to the fact that the majority of studies were either conducted in the USA, in Germany, or in Japan. As a consequence, the present review cannot draw conclusions about the influence of people's culture on their beliefs about social robots. Additionally, we would note that the geographical location in which the studies were conducted is only an approximation of participants' cultural background as most studies did not report this information. Even where the nationality and/or ethnicity of participants was reported, it may not necessarily reflect the participants' cultural background. The present review identified only six studies labelled as cross-cultural which may indicate a lack of cross-cultural research on people's attitudes toward social robots.

Similarly, our data and findings do not provide a strong enough base for conclusions regarding the extent to which the design (i.e., level of human-likeness) and application area of the robot moderated people's attitudes and anxiety toward, trust in, and acceptance of robots. We also found no evidence that the age of participants was associated with any of the outcomes despite existing empirical evidence to the contrary. Previous studies comparing young and elderly adults have demonstrated that, in general, older adults have more negative attitudes toward robots and are less willing to use robotic technology (Wullenkord, & Eyssel, 2014). We did find evidence that the gender of participants was associated with the extent to which they trusted robots (in general, samples with a larger percentage of female participants reported more trust in robots). However, for most outcomes, the number of studies was quite small and it was difficult to draw clear conclusions regarding the effect of gender.

3.7.4. Have attitudes changed over time?

We found no evidence that beliefs about social robots have changed over time. However, the earliest paper in our review was published in 2005 and the majority of studies were published between 2014 and 2018. As such, our analysis was based on a rather constrained data set with the majority of data points falling within a four-year period. While we cannot say for certain whether people's beliefs about social robots have changed over time, we should probably first ask whether social robotics has existed long enough for such changes to have occurred at all.

One approach might be to consider the changes in attitudes, trust, and acceptance that have taken place in relation to robotics in general and past technological developments such as the modern computer and smartphones and then use these trends to predict how peoples' beliefs about social robots might change over time. For example, Gnambs and Appel (2019) investigated changes in attitudes towards robotic systems within the European Union between 2012 and 2017. They found that, although attitudes toward various robotic systems were generally positive, there was a significant decrease in favourable opinions over the five-year period. Most notably, attitudes towards autonomous robots in the workplace were overall the most positive but also saw the largest negative shift in attitudes between 2012 and 2017. Gnambs and Appel proposed that the change in people's attitudes may be the result of increasing media coverage of robotic systems and growing fears about automation and its impact on the job market (Broadbent et al., 2012; Ebel, 1986). Therefore, although the present review suggests that people's attitudes toward social robots are typically slightly positive, it may be that we should expect a negative shift in attitudes over the coming years.

3.7.5. Suggestions for future research

The present review identifies a number of methodological issues that should be addressed by future research. Some of these limitations are not specific to the study of social robotics – for instance, the tendency to rely on samples of student volunteers. Although practical and financial limitations are often a barrier to the acquisition of more diverse sample groups, it is important to acknowledge the limitations of sampling procedures and consider potential bias when drawing conclusions. Where broader questions about the way that robots should be designed and integrated into specific domains are asked, it is important to acknowledge that making broader generalisations about the rest of society based on this limited sample of participants may not be appropriate. A further observation was the reliance on self-report measures (typically multi-item Likert scales). While using self-report measures often makes sense and yields useful data, some consideration should be given to applying other types of measures alongside well-known scales such as the NARS, especially given the intention-behaviour gap in technology usage (Bhattacherjee, & Sanford, 2009). Indeed, there have been advances in both behavioural and / or physiological measures (e.g., of arousal) that may prove useful in future research.

Finally, we attempted to analyse the effect of previous experience with robots on participants' attitudes, as research has found that this might play a role in shaping people's beliefs about robots (Kachouie et al., 2014; Leite et al., 2013; Syrdal et al., 2014). Although fourteen studies reported information about the extent to which participants had interacted with social robots previously, there was only one study in which more than half of the participants had seen or interacted with a robot before. The rest of the studies reported that the majority of participants had little to no experience with social robots. As such, the findings of our review should probably be considered a reflection of people's initial attitudes toward social robots; something that – given that most people rarely have any contact with social robots – is likely to currently reflect most people's attitudes toward social robots. Readers interested in the effect of long-term interactions on attitudes might consult a review by Leite et al. (2013), which suggests that, while people are generally willing to interact with robots repeatedly, their attitudes may change over time.

3.8. Conclusion

The evidence presented in this review suggests that people – at least people who do not have extensive experience of social robots - generally have a positive view of social robots.

More specifically, the evidence suggests that people typically have positive feelings and thoughts toward social robots and are willing to interact with robots should the chance present itself. These findings may help to alleviate some of the concerns regarding the likelihood that people will adopt robotics in socially focused domains such as healthcare and education. However, knowing that people typically have somewhat positive beliefs about social robots does not necessarily help us to predict the economic and social impacts of widely adopting this type of technology. A positive disposition is only one of a number of factors that may determine the landscape of human-robot relationships in the future and we suggest that applying theories of inter-group relations and attitude-behaviour models (Pettigrew, 1998; Pratto et al., 2006; Hewstone, & Swart, 2011) to the study of social robotics might help to understand what these relationships may look like. Finally, although we may draw parallels between the progression and impact of other technology (such as computers) and social robotics, we should also acknowledge the qualities that mark social robots as not just another technological development but perhaps as an entire new social group with its own complexity (Prescott, 2017).

Chapter IV

Study 2: The effect of direct and extended contact on trust in and attitudes towards social robots

4.1. Overview

This chapter is the first empirical study of the thesis. It investigates how direct and extended contact with robots affects people attitudes towards robots. The research protocol for direct contact consisted on participants interacting with Pepper (a humanoid robot) while the robot was giving information about hair care products. After that, participants recorded a video that had to be seen by another participant, which would be in the extended contact condition. Two control conditions were also implemented, one of them in which participants had extended contact with another human being, and another one in which participants had no contact with anyone or any robot. Trust, explicit and implicit attitudes were measured by the means of two questionnaires and an IAT. However, while performing the study, there were some issues (for example, participants not showing up) and confounding variables, especially in the extended contact condition, that could have affected the outcomes.

4.2. Introduction

The aim of the present research was to examine the effects of direct and extended contact on people's attitudes, both implicit and explicit, towards robots. Our hypothesis was that attitude formation and change with respect to robots has similar dynamics to attitude change with respect to interpersonal relations; particular, attitudes towards members of minority groups with whom people rarely have contact. In other words, based on the extension of Allport's (1954) contact hypothesis to human-robot interaction, it was predicted that both direct and extended contact experiences with social humanoid robots can influence attitudes towards this technology.

In the present research, social psychology techniques were borrowed and used in a HRI setting. The media equation (Nass, & Yen, 2010; Reeves, & Nass, 1996) suggests that people are able to treat communication media and artificial intelligence as if they were real people or social agents. Some experts say that this happens because people anthropomorphise this type of technology (Bartneck, et al., 2007b) while other claim that anthropomorphism is not always necessary for the media equation effect to happen (Nass, & Moon, 2000). This phenomenon is also present when people interact with social robots (Reuten, van Dam, & Naber, 2018). Because people are able to see social robots as social agents, in the field of HRI, it is common to use social psychology techniques and adapt them to be used in settings where participants interact with robots. The present research borrowed intergroup research methods, which were tailored to this specific study.

While there are some studies that investigate how direct and imagined contact with a robot can affect people's attitudes towards robots (e.g., Kuchenbrandt & Eyssel, 2012), no research to date has specifically investigated the effects of extended forms of contact on attitudes and trust towards robots. However, taking into account all the information given in the previous chapters, it could be hypothesised that extended contact can also influence attitudes towards robots. Specifically, if a person knows that a friend or relative had a positive or negative experience with a humanoid robot, then this could influence his or her attitudes towards robots. Therefore, the aim of this research is to find out if indirect contact has an effect on people's attitudes towards humanoid robots and to compare these effects with that of direct contact. In this study particularly, the type of indirect contact that was used is a mix between extended and mediated contact. Participants having indirect contact with the robot watched a video of someone they know talking about the robot. It is extended contact because the person in the video had met the robot and then talked about it to their friend.

This research can help us to know if different types of contact are effective for people to see robots from a realistic perspective and, by seeing the differences and similarities between different contacts, this method could be applied in the future to help people base their attitudes and trust on realistic references. That is to say, these types of contact could be used as a tool to help people to acquire knowledge about robots.

4.3. Robots used in this thesis

Several humanoid robots have been designed to interact with people. This section describes particularly Pepper and NAO since those are the robots used in the empirical studies in this thesis (Pepper was used in Studies 2 and 3, while NAO was used in Study 4). Since these studies are focused on humanoid robots, there was a need to have two robots that matched this characteristic while being different enough in order to reduce the possibility of having results linked to a specific robot and provide a conceptual replication of the effect of different forms of contact on attitudes.

Figure 9: The humanoid robot NAO next to a person.

NAO (Figure 9) is the first humanoid robot created by Aldebaran robotics, which was then acquired by SoftBank Group in 2015 and rebranded as SoftBank Robotics (Crunchbase, 2019). NAO's first prototype was developed in 2008. According to Softbank Robotics (2019) the last generation of NAOs (V6) measure 574 millimetres and weight 5.5 kilograms. It has a head, two arms, torso, and two legs. It can be purchased with a suitcase; together with its size and weight, this makes NAO very easy to transport and therefore, it can be used in many different environments. Its popularity can be associated to NAO's "wide availability, appealing appearance, accessible price point, technical robustness, and ease of programming" (Belpaeme, et al., 2018, p. 4). NAO is able to recognize voice and process language. The speakers allow the robot to speak and reproduce text-to-speech strings. It uses the cameras for computer vision and facial recognition. It has 25 degrees of freedom, which enable the robot to move and adapt to its environment. NAO has four microphones, sonar rangefinder, two infrared emitters and receivers, inertial board, nine tactile sensors and eight pressure sensors. It also comes with an Ethernet port and Wi-Fi. It also comes with a software called Choregraphe that can be used to program the robots in a drag and drop setting. Moreover, NAO is open and fully programmable, which means that, apart from using Choregraphe, developers can also program the robot in C++, Python, Java, MATLAB, Urbi, C and .Net. It also comes with a simulation software package so users can code their program and then run simulations without the need of using NAO physically.

NAO has been frequently used in HRI studies (e.g. Alemi et al., 2014; Conti et al., 2017; Fridin, & Belokopytov, 2014; Hosseini et al., 2017; Kim et al., 2016; Kuchenbrandt, & Eyssel, 2012; Mirnig, Stadler, Stollnberger, Giuliani, & Tscheligi, 2016; Stadler, Weiss, & Tscheligi, 2014; Stanton, & Stevens, 2017; Wullenkord, & Eyssel, 2014; Wullenkord et al., 2016). NAO has also become a significant tool in education. It has been used in schools to introduce robotics and the new technologies to children (Active-robots, 2014; Belpaeme et al., 2018; Mubin, Stevens, Shahid, Al Mahmud, & Dong, 2013). It has been used to help autistic children in the classroom since they perceive that it is less threating to interact with a robot than a person because robots have no emotions and are not judgemental (Burns, 2012). Research with robots and autistic children is not new and NAO has played a big role helping researchers carry out their investigations in this field (Feng, Gutierrez, Zhang, & Mahoor, 2013; Greczek, Kaszubski, Atrash, & Matarić, 2014; Miskam et al., 2013; Shamsuddin et al., 2012a; Shamsuddin et al., 2012b; Taheri, Alemi, Meghdari, Pouretemad, & Holderread, 2015; Tapus et al., 2012). Apart from this, NAO has also been used as an assistant to welcome, inform and entertain visitors in healthcare centres and companies (New Atlas, 2015; Marous, 2015; The Guardian, 2015).

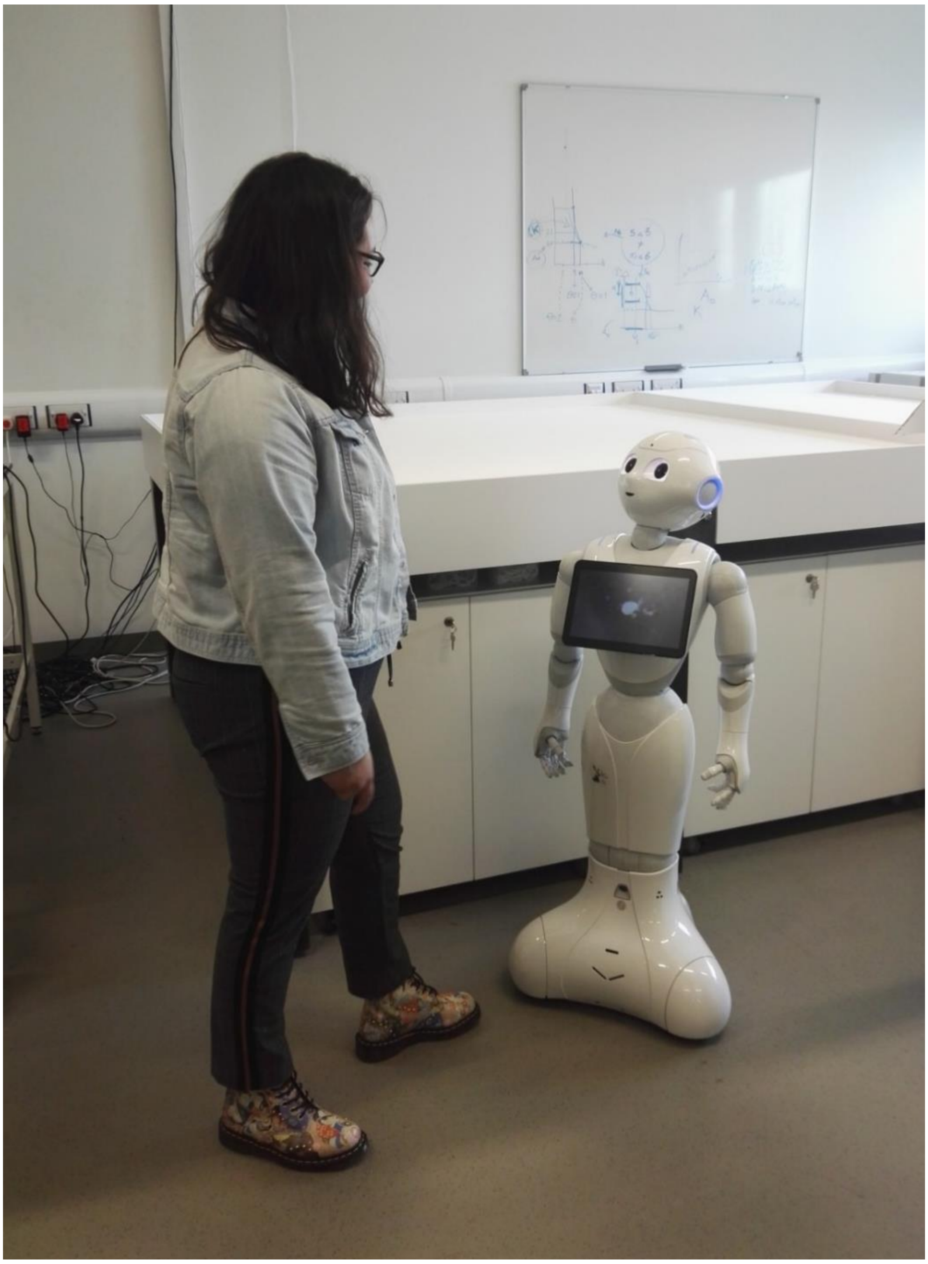


Figure 10: The humanoid robot Pepper next to a person.

Pepper (Figure 10) is another humanoid robot developed by Softbank robotics with the first prototype created in 2014. The robot has a head, two arms, a torso and one leg with omnidirectional wheels. It also has a tactile tablet on its torso that can be used to access the settings, show images, videos or applications. Compared to NAO, Pepper is tall (120 centimetres) and heavy (28 kilograms). This makes it more difficult to transport than NAO. In addition, Pepper is more expensive. Pepper's physical traits are not arbitrary and were established after meticulous research. The developers who designed Pepper's appearance took the 'uncanny valley' hypothesis into account and made the robot in a way that has a human appearance without falling into the valley (Pandey, & Gelin, 2018, Mori, 1970). In addition, its shape was aimed to be gender neutral to avoid gender stereotypes (Pandey, & Gelin, 2018). Pandey and Gelin also explain that Pepper's height was decided based on empirical evidence suggesting that a daily-life robot should be taller than NAO but shorter than an average person sitting on a chair, hence the 120 centimetres. In addition, its curvy figure makes the robot appear safe and friendly (Softbank Robotics, 2019). According to Softbank Robotics, Pepper has 20 degrees of freedom which enables the robot to have an expressive body language. It also has 4 directional microphones and a natural language processing system that allows responsive dialogue options. It uses two 2D cameras and a 3D sensor. This grants Pepper with facial recognition capabilities, which can be used to detect people's emotions as well as identify different users and make interactions more interesting. It also has tactile sensors in its head and hands so developers can program Pepper to react to touch. The tablet makes it easy to access content to highlight messages and support speech. Pepper has a six-axis inertial measurement unit (IMU) composed of a three-axis gyrometer with an angular speed of ~500 °/s and a three-axis accelerometer with an acceleration of ~2 g. Inside the inertial board, an algorithm is implemented to compute the base angle from the accelerometer and gyrometer. This inertial unit measures Pepper's attitude or orientation (yaw, pitch, and roll), speed and position (Pandey, & Gelin, 2018). This is important for the robot to navigate and know its position. In addition, Pepper's base is equipped with three laser sensors, two sonar sensors, two infrared sensors, 6 laser actuators and 3 bumpers. All of these sensors help Pepper to detect obstacles and protect the robot's integrity as well as the people close to the robot. It also has a Wi-Fi connection and an Ethernet port. Like NAO, Pepper comes with the software Choregraphe and the simulation software. It is also open and fully programmable in Python, Java, C++ and Android.

Pepper was initially designed to be a customer service robot working in Softbank stores. However, users became more and more interested in the robot and its application expanded to other areas such as welcoming visitors and guiding customers in shops and commercial areas (Pandey, & Gelin, 2018, New Atlas, 2015; Marous, 2015). As people's interest awakened, the use of the robot extended to other areas like research (Foster et al., 2016; Qureshi, Nakamura, Yoshikawa, & Ishiguro, 2016) or education (Belpaeme, et al., 2018; Tanaka et al., 2015). Nonetheless, in spite of Pepper's advantages, it could be possible that Pepper's size and price (compared to other robots like NAO) is affecting its popularity in research and education.

Pepper and NAO are similar in the sense that they are both humanoids and have a friendly appearance but their size is significantly different, with Nao being smaller. In addition, there are also differences in shape such as the base of the robot: NAO is bipedal while Pepper only has one leg. They also have different hands; NAO has 3 fingers in each hand while Pepper has 5 fingers. Apart from that, their overall appearance also differs in the sense that NAO has more angular and squared shapes while Pepper is more round and has more curves. In addition, Pepper's eyes are bigger than NAO's in proportion to its head, which gives Pepper a childish friendly face.

The three empirical studies presented in this thesis were very similar in the sense that they aimed to investigate the effect of direct and extended contact on people's attitudes towards social humanoid robots. However, the use of different robots allowed this research to have results that were not tied directly to a particular robot.

4.4. Research Hypothesis

Both direct and extended contact with robots will affect people's attitudes towards robots. That is to say, there will be a main effect of contact in trust, explicit and implicit attitudes towards robots. There will also be an interaction with time and condition. Participants will have more positive attitudes after having wither direct or extended contact with the robot.

4.5. Method

4.5.1. Participants

Forty-six participants took part in the experiment, 25 males and 21 females. Their mean age was 24.63 (SD = 6.11), most of them were British (N = 26, 56.5%) and they were all students and staff from a university in the north of England. Three methods of recruitment were used. First, the online SONA system was used, which gives course credits to psychology students who take part in research studies. In this case, participants who used this method received 4 credits for their participation. The second method was the nominations system; some participants had to nominate two friends who would receive an invitation to come to perform the experiment later or another day. During the recruitment process, there were some issues such as participants not showing up and, therefore, another recruitment method was implemented. The third recruitment method was a mailing list which sends an invitation to all students and staff in the university. Participant who did not use the online SONA system did not receive any incentive. Participants came from several different backgrounds such as psychology (N = 8), robotics (N = 6), natural sciences (N = 5), healthcare (N = 5), engineering (N = 5), languages (N = 4), media studies (N = 3), business (N = 2), librarianship (N = 2), archaeology (N = 1), charity (N = 1), gardening (N = 1), philosophy (N = 1), politics (N = 1) and dual honours (French/music) (N = 1).

The issues recruiting participants created a domino effect influencing the distributions of participants between conditions, which, at the same time, could have also affected the results. This concerns are expanded in the discussion section of this study.

4.5.2. Measures

There are several standardized questionnaires that are designed to measure people's attitudes towards, robots. One of them is the **Negative Attitudes towards Robots Scale (NARS**) (Nomura, et al., 2004). It was developed for measuring peoples' attitudes towards communication robots in daily-life. It is based on a Likert scale from 1 to 5 where 1 means that the participant is not anxious at all and 5 shows the highest grade of anxiety. Its internal consistency, factorial validity and test reliability have been tested and confirmed repeatedly (Bartneck, et al., 2005b; Bartneck, et al., 2007a; Cramer et al., 2009; Katz & Halpern, 2014; Nomura et al., 2006b; Nomura, et al., 2007; Syrdal, Dautenhahn, Koay, & Walters, 2009). Some of the questions in NARS are the following:

I would feel uneasy if I was given a job where I had to use robots.

I would feel paranoid talking to a robot.

I feel that if I depend on robots too much, something bad might happen.

I feel comforted being with robots that have emotions.

The values in the NARS go from 1 to 5. In order to compute the score for each participant, items 3, 5 and 6 need to be reversed following Nomura's instructions. That is to say, if the participant's answer was "1", it needed to be reversed to "5"; if the participant's answer was "2", then it was reversed to "4", and if the answer was "3" (neutral), it was kept the same. In the following formula, the reversed score is represented by *R*, while the original answer from the participant is represented by *a*.

$$R = (5 - a) + 1$$

After that, all the 14 answers of the questionnaire were summed and the average was computed following this formula:

$$\bar{x} = \frac{1}{14}\sum_{i=1}^{14} x_i$$

There is empirical evidence that trust influences acceptance of new technologies (Eiser et al., 2002). Hancock et al. (2011b) define trust "as the reliance by an agent that actions prejudicial to their well-being will not be undertaken by influential others." (p. 24) Another definition is given by Freedy et al. (2007); trust is "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability." (p. 108) There is another definition that addresses the importance of vulnerability. It is the one given by Mayer, Davis and Schoorman (1995), who define trust as "the willingness of a party to be vulnerable to the outcomes of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party." (p. 709) People determined their trust in robots by observing how robots accomplish the individual's goals and the manner in which this process is transparent (Freedy et al., 2007). Therefore, a person who trusts a robot will think that it is capable to perform its task successfully, following the instructions and without harming anyone. If that is the case, this would have a positive influence on the robot's acceptance and, consequently, a positive impact on human-robot interaction (HRI) (Hancock et al., 2011b).

In order to measure trust in robots, the proposed research will use an adaptation of the **40 item human-robot trust scale**, developed by Schaefer (2013). This scale has 40 questions and was developed to measure trust specifically in a HRI environment. The scores are percentages where 0 shows no trust at all and 100 is the highest level of trust. It was designed to consider multiple forms of trust like cognitive trust, affective trust, and trustworthiness. Moreover, this scale can be used by any robotic domain from industry to military, to the everyday robot. This scale has been tested and validated by a set of 6 studies carried on by their developers and it has also been used in other studies since then (Kessler, et al., 2017; Volante, et al., 2016). So, for all the reasons mentioned previously, it fits the purpose of this study. It presents questions such as:

What % of the time do you think this robot will act consistently?

What % of the time do you think this robot will protect people?

What % of the time do you think this robot will act as part of a team?

What % of the time do you think this robot will function successfully?

Since the questions refer to "this robot", it would not make sense to use it before participants meet the robot. In addition, it would not be useful in our control conditions because participants will not be interacting with any robot in these conditions. Thus, the questions were adapted and instead of "this robot", they are going to say "robots". Therefore, the questions will be such as the following:

What % of the time do you think robots will act consistently?

What % of the time do you think robots will protect people?

What % of the time do you think robots will act as part of a team?

What % of the time do you think robots will function successfully?

In order to calculate the scores for the Trust scale, a process similar to the one implemented for the NARS was followed but this time, for each answer, the values went from 0 to 100. There were 5 items that needed to be reversed (5, 7, 9, 29, 34) following this formula:

$$R = (100 - a)$$

Then, all the 40 answers were summed and the average was computed using the following formula:

$$\bar{x} = \frac{1}{40}\sum_{i=1}^{40} x_i$$

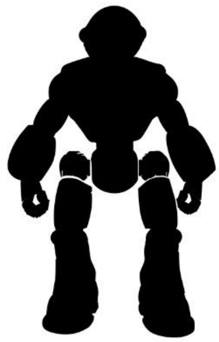
In order to test the internal reliability of these questionnaires, a Cronbach's alpha test was run on each on them. The NARS (14 items) that participants made before the experiment has a score of .541, which indicates a poor internal consistency. The NARS after the experiment scored a value of .847, which suggests a good level of internal reliability. The Trust scales (40 items) that participants made before and after the scored a value of .898 and .922 respectively, which indicated a good level of internal consistency.

In order to measure implicit attitudes, this study implemented an IAT adapted from MacDorman et al. (2009), which evaluates the association between two target concepts, in this case humans and robots, and two attributes, which in this case are pleasant and unpleasant words. In the present research, when participants completed the IAT, they first sat in front of a computer and they were instructed to classify certain silhouettes (humans or robots) and words (pleasant and unpleasant) accordingly by either pressing the response key “A” (left) or “L” (right) on the keyboard. The instructions were on the screen and also explained until the participants understood what they had to do.

When the trials had begun, participants saw at the centre of screen either a silhouette (a robot or a human). At the top of the screen, they could see the words “Robot” and “Human”, on either side of the screen; then they had to press either “A” (left) or “L” (right) in order to classify the silhouettes accordingly. After that, they saw words at the centre of the screen (a pleasant or an unpleasant word). On either side at the top of the scree, they could read “Pleasant” and “Unpleasant” and they just had to follow the same process pressing either “A” or “L”. After that, they had to do the same but this time human silhouettes and pleasant words had the same response key, and robot silhouettes and unpleasant words also shared the same response key. For example, if they saw on the top right “Robot or unpleasant” and on the top left “Human or pleasant”, they had to press “L” (right) every time they saw a robot silhouette or an unpleasant word at the centre of the screen, and “A” (left) when they saw a human silhouette or a pleasant word. Once they had press the key, the next trial appeared on the screen automatically. If they made a mistake, a red cross would appear at the centre of the screen until they had press the right key (“A” or “L”). Each block had a certain amount of trials (Table 8) and once the block was finished, participants were able to read the instructions again and continue with the following block. Table 7 shows some examples of the trials used in the actual study.

Table 6

IAT trial examples

Example of a trial in the IAT	Response key
Human or pleasant Robot or unpleasant 	Right (L)
Human or pleasant Robot or unpleasant Terrible	Right (L)
Human or pleasant Robot or unpleasant 	Left (A)

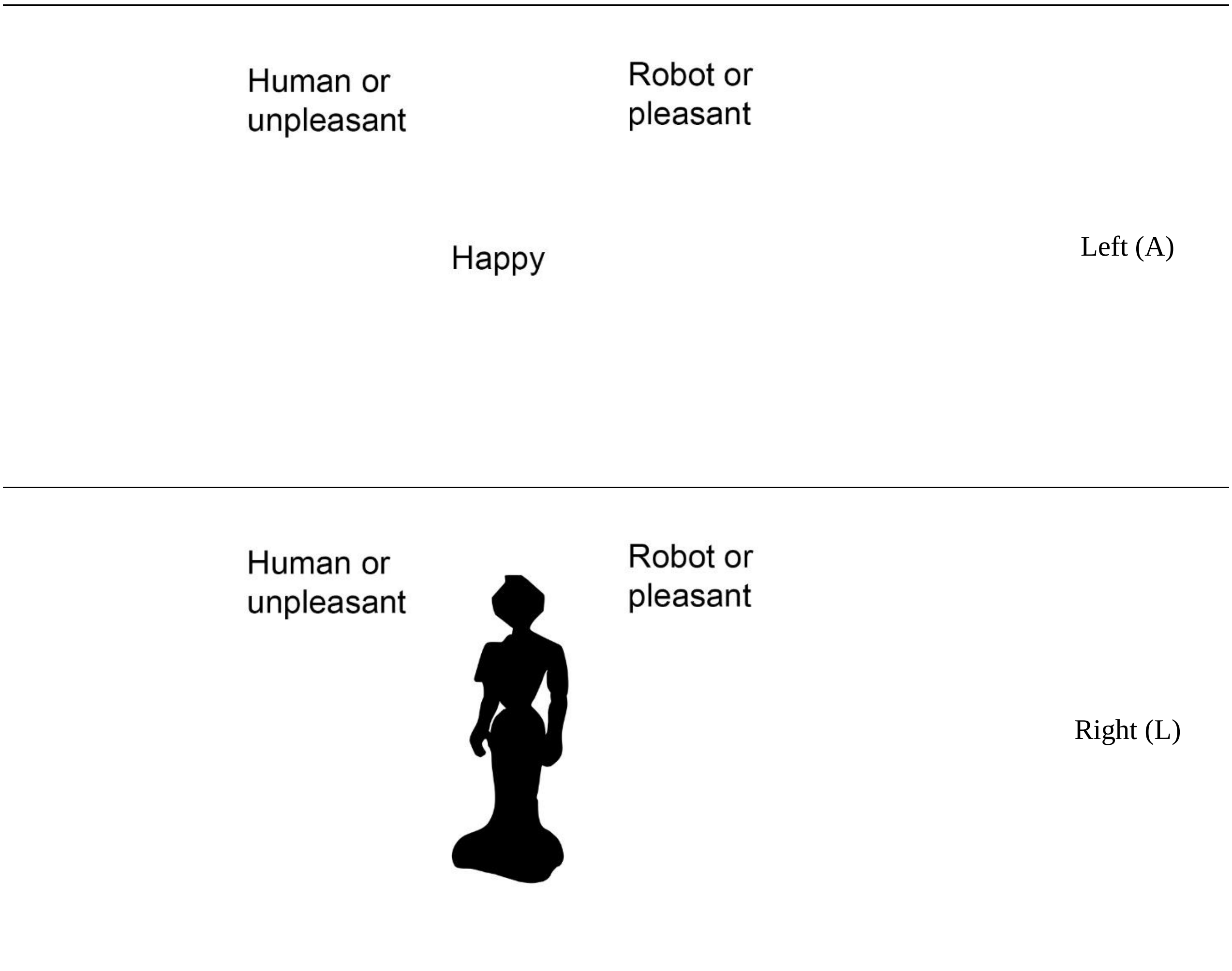


Table 8 shows the seven steps (blocks) of the IAT procedure. Blocks B1, B2 and B5 were there so the participants got used to the task. Blocks B3 and B4 are congruent blocks because stereotypically associated word valences and robot/human categories share a response key (human/pleasant and robot/unpleasant). Blocks B6 and B7 are incongruent because robot/human categories share the response key with non-stereotypic word valences (human/unpleasant and robot/pleasant). The IAT measures the reaction time between the stimuli (silhouette or word) appearing on the screen and the participant pressing the correct response key. Participants were instructed to perform the task as quickly as they could and everyone completed the IAT in the laboratory using the same computer. A faster reaction in the congruent blocks would mean that the participant associates humans to pleasant words and robots to unpleasant words. A faster reaction in the congruent blocks would portray a stronger association between robots and pleasant words, and humans and unpleasant words.

By using this method, it is expected that the extracted data will be more objective than the data given by the explicit methods because IATs do not rely on introspection and, therefore, it is more difficult to give a biased answer.

Table 8
Sequence of blocks in the IAT

Block	Number of trials	Items assigned to left-key response	Items assigned to right-key response
B1	20	Human	Robot
B2	20	Pleasant	Unpleasant
B3	20	Human or pleasant	Robot or unpleasant
B4	40	Human or pleasant	Robot or unpleasant
B5	40	Robot	Human
B6	20	Robot or pleasant	Human or unpleasant
B7	40	Robot or pleasant	Human or unpleasant

For the IAT, the D score was calculated by using an algorithm described by Greenwald et al. (2003); the higher the D score, the more negative are attitudes towards robots. All participants performed the IAT once they had arrived using the same computer. The control group only did the IAT once because they had no contact with the robot and therefore, there was no before or after contact. In order to compute the D score, data from the blocks B3, B4, B6 and B7 were used. Extreme-value and error latencies were removed or treated according to Greenwald et al. (2003). For each of the four blocks, the resulting values were averaged. Then, one pooled standard deviation was calculated for all trials in B3 and B6 (SD_1); and another for B4 and B7 (SD_2). The D score was computed using the following formula in which B3, B4, B6 and B7 represent the average value for each block.

$$D = \frac{\left(\frac{B6 - B3}{SD_1} + \frac{B7 - B4}{SD_2}\right)}{2}$$

4.5.3. Robot programming

Pepper was programmed to detect keywords while talking to the participants by using the feature “choice” in the integrated software Choregraphe. First, the robot introduced himself to the participant and then it started asking questions about their type of hair and recommend hair care products. The system was programmed to detect keywords that were pertinent to the question asked. For a full list of all the keywords, see Supplementary Material 4. These keywords could be integrated in a sentence. For example, if a participant said “I have long hair”, that was enough for the system to detect the keyword “long hair”. Participants were recorded during the interaction. The most common issue was participants having to repeat their sentences or words so the robot could process them. This depended on the accent and tone of the participants, which were instructed to speak loud and clear, and repeat their sentences if the robot was not responding. However, in general, interactions were smooth and participants were able to have a conversation about shampoo.

4.5.4. Procedure

Ethical approval was obtained by Department of Psychology Research Ethics Committee. In order to test the hypotheses, there were 2 experimental groups of participants and 2 control groups. All participants completed the NARS and Trust scale online before they came to the research centre. Implicit attitudes towards robots were assessed once they had come to the centre before they took part in the experiment and all participants used the same computer to do that. All the measures were taken again once they had performed the experiment.

The first experimental group was the **direct contact (DC) condition**, in which participants were given the opportunity to interact directly with the humanoid robot Pepper. Participants were given instructions to interact with the robot pretending that they were in a haircare shop and Pepper was the shop assistant (the shop contains a range of haircare products). The conversation was limited to 5 minutes. Once they were finished, they completed the measures of their beliefs about robots a second time. After that, participants in this condition were asked to record two short video messages (between 1:30 min and 1:40 min long) that participants in other conditions would watch. In the first video, they were asked to describe their interaction with the robot. For this first video, participants were given the following written instructions:

> *You are going to record a video for a friend. You will explain your experience with Pepper. Please answer the following questions while talking to the camera:*

1. *What happened since you first saw Pepper until you left the room?*
2. *What did you talk about?*
3. *How did you feel interacting with Pepper?*
4. *Did Pepper help you achieve the purpose of the conversation (deciding which product you like)?*
5. *Did you like Pepper? If so, why? If not, why not?*

Then, they had some time to think about the answers and write some notes if they felt that they needed to. After that, they recorded the video talking to the camera.

In the second video, they were asked to describe a recent conversation that they had had with someone new, who they had met for the first time relatively recently. This task was chosen to ensure that both interactions described in the videos (i.e., an interaction between a human and a robot and between a human and a human interaction) would be more similar – i.e., they would both involve a novel interaction. For this video, participants were given the following instructions:

> *You are going to record another video describing a conversation that you had recently with someone you met recently for the first time; it can be a conversation with a new friend, a shop assistant, a waiter, a taxi driver, a receptionist or anyone you have meet recently. Please answer the following questions while talking to the camera:*

1. *What happened with this person since you began this conversation until you finished talking?*
2. *What did you talk about?*
3. *How did you feel interacting with this person?*
4. *If the conversation had a purpose did this person help you achieve the purpose of the conversation?*
5. *Did you like this person? If so, why? If not, why not?*

Then, they had some time to think about the answers and write some notes if they felt that they needed to. After that, they recorded the video talking to the camera.

Once these participants finished recording these two videos, they were debriefed and they were asked to nominate two friends that, allegedly, would come to perform the experiment, be placed in the other conditions with extended contact and watch the videos. This was done because, by definition, extended contact with robots is knowing someone who has interacted with robots previously. In other words, for extended contact to be effective, it is essential that the person receiving the extended contact perceives the person giving the contact as an in-group member. Therefore, it is relevant that the participant in any of the extended contact conditions is familiar with the person in the video. By using this nomination system, it was ensured that these two participants knew each other. The contact information of these nominated friends were collected and they were invited to take part in the experiment as a participant.

The second experimental group was the **extended contact (EC) condition**, in which participants were asked to watch the video that their friend in the DC condition had made, talking about his or her experience with the robot. They received the following written instructions:

> *You are going to watch a video of a friend talking about his or her social interaction with the humanoid robot Pepper. Pay attention to the video and, after that, you will be given some questionnaires.*

After that, they completed the measures of their perception about robots a second time.

The first control group was the **extended contact control (ECC) condition**; in which participants were asked to watch the video that their friend in the DC condition had made talking about a conversation that they had had with someone new that they met only recently.

> *You are going to watch a video of a friend talking about his or her social interaction with someone else. Pay attention to the video and, after that, you will be given some questionnaires.*

After that, they completed the measures of their perception about robots a second time.

Participants in the second **control condition** had no contact with the robot. They first completed the explicit measures of their perception about robots at baseline and they completed them again once they came to the laboratory. As stated before, all participants used the same computer in the laboratory to complete the IAT. They could not have done the IAT before coming to perform the experiment. So, participants in this control condition only completed the IAT once because, taking into account that they did not interact with anything, it would not have made sense for them to complete the IAT twice in a row.

4.6. Results

Skewness and kurtosis were used in order to test the normality of the data. All the values fell between -1.96 and +1.96, which is considered acceptable in order to prove normal univariate distribution (George, 2011). Table 9 presents the descriptive statistics for the key measures by condition. The histograms for each dependant variable in each condition are presented in Supplementary Material 4.

4.6.1. Randomization check and correlations between measures

A series of one way independent ANOVAs were conducted in order to check that the attitudes of participants in the four conditions did not differ at baseline. There were no significant differences in explicit attitudes, $F(3, 42) = 1.64$, $p = .194$, partial eta^2 = .11, trust, $F(3, 42) = 1.88$, $p = .148$, partial eta^2 =.12, or implicit attitudes, $F(3, 42) = 0.70$, $p = .558$, partial eta^2 = .05, between the conditions at baseline. There were no correlations between any measure.

4.6.2. Effect of direct and indirect contact on attitudes

In order to examine the effects of direct and indirect contact (relative to control procedures) on attitudes towards robots and trust, we conducted two 4-between (condition: direct contact, extended contact, extended contact control, control) by 2-within (time: before vs. after) mixed ANOVAs with explicit attitudes and trust as the dependent variables. This was done primarily to find an interaction between condition and time although main effects are also reported.

The main effect of condition on explicit attitudes towards robots was not significant, $F(3, 42) = 1.65$, $p = .19$, partial eta^2 = .11. There was not a significant effect of time, $F(1, 42) = 3.42$, $p = .07$, partial eta^2 = .08; and there was not a significant interaction between condition and time, $F(3, 42) = 1.26$, $p = .30$, partial eta^2 = .08.

The main effect of condition on trust towards robots was not significant, $F(3, 42) = 2.50$, $p = .07$, partial eta^2 = .15. There was not a significant effect of time, $F(1, 42) = .33$, $p = .57$, partial eta^2 = .01; and there was not a significant interaction between condition and time, $F(3, 42) = .59$, $p = .63$, partial eta^2 = .04.

A one-way ANOVA was conducted to examine the effect of condition on implicit attitudes following the procedures.[1] There was not a significant main effect of condition on implicit attitudes, $F(2, 31) = .69$, $p = .509$, partial eta^2 = .04.

---

[1] It was not possible to run a repeated measures ANOVA as implicit attitudes were only measured once in one of the control conditions.

Table 9

Study 2: Descriptive statistics

		Direct contact						Extended contact					
		Explicit attitudes		Trust		Implicit attitudes		Explicit attitudes		Trust		Implicit attitudes	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Mean	2.32	2.38	61.24	62.38	0.41	0.20	2.49	2.58	70.04	73.53	0.44	0.35
	SD	0.55	0.37	10.42	14.15	0.33	0.45	0.51	0.24	10.29	10.41	0.33	0.50
Skewness	Statistic	0.08	-0.06	-0.84	-1.91	-0.71	-0.03	-0.12	0.69	-0.01	-0.48	0.06	-0.28
	Std. Error	0.64	0.64	0.64	0.64	0.64	0.64	0.66	0.66	0.66	0.66	0.66	0.66
Kurtosis	Statistic	-1.31	-0.63	0.22	4.53	-0.70	-0.95	-0.63	0.31	-1.48	0.41	-1.19	-1.20
	Std. Error	1.23	1.23	1.23	1.23	1.23	1.23	1.28	1.28	1.28	1.28	1.28	1.28

Table 9 (continuation)

Study 2: Descriptive statistics

		Extended contact control						Control					
		Explicit attitudes		Trust		Implicit attitudes		Explicit attitudes		Trust		Implicit attitudes	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Mean	2.71	2.69	59.53	59.04	0.48	0.14	2.23	2.53	62.79	61.68	0.56	-
	SD	0.46	0.37	9.53	7.87	0.24	0.33	0.65	0.45	13.95	18.43	0.13	-
Skewness	Statistic	0.51	1.05	0.06	0.19	0.54	-1.58	0.59	-0.04	-0.79	-0.56	-0.25	-
	Std. Error	0.66	0.66	0.66	0.66	0.66	0.66	0.64	0.64	0.64	0.64	0.64	-
Kurtosis	Statistic	-0.52	0.95	-1.45	-0.29	-0.01	3.37	-0.06	-1.16	1.89	0.30	-0.29	-
	Std. Error	1.28	1.28	1.28	1.28	1.28	1.28	1.23	1.23	1.23	1.23	1.23	-

4.7. Discussion

The results obtained suggest that neither direct or extended contact had any effect on participants' attitudes or trust in robots. This contradicts previous studies in which direct contact and other types of indirect contact with a robot actually changed participants' attitudes (Conti, Di Nuovo, Buono, & Di Nuovo, 2017; Wullenkord, & Eyssel, 2014; Zlotowski et al., 2015). It also contradicts the contact hypothesis (Allport, 1954) since any of the types of contact had an effect.

However, these results could have been affected by several limitations that the study had. The most relevant concern was that the friends that participants nominated did not respond to the invitation and so participants the extended contact was not with someone who the participants knew. So, while performing the study, the method of recruitment had to be changed in order to recruit participants to the extended contact conditions. Therefore, two recruitment methods were used: (i) the SONA system (which was used since the beginning of the experiment) and (ii) a mailing list to all students and staff in the university.

This situation created a domino effect and other elements of the study were affected. At first, when the online SONA system was used, participants were being allocated in the DC condition and the control condition, while we were waiting for the nominated friends to come (to take part in the EC condition and the ECC condition). However, the nominated friends never came. The result was that, at the beginning of the experiment, all participants were in the DC condition and the control condition, while EC and ECC conditions were empty. It was therefore decided to abandon nominations and instead simply to recruit participants using a mailing list. Therefore, participants in the EC and ECC conditions did not know the person who they were watching in the video, which could have affected how extended contact was implemented. This in addition to the non-random allocation (i.e., recruiting DC and control participants first) may have compromised the validity of the data.

In addition, a power analysis was not conducted before the study. So, it could be possible that the sample size was too small and, therefore, a type II error (i.e. a false negative) could be happening due to insufficient power. This needs to be addressed in the following studies by running a power analysis in order to know how many participants should take part in the study.

4.8. Conclusions

The aim of this Study 2 was to investigate the effects of direct and extended contact on people's trust and attitudes towards robots. In order to do that, participants were exposed face to face to a robot, watched a video of someone talking about the robot or were allocated in any of the control conditions. The results indicate that none of these procedures had any effect on the dependent variables. However, Study 2 has some limitations that need to be discussed. First, the fact that participants in the EC condition did not know the person in the video might have affected the results and maybe if they had known the person they were watching, the impact would have been higher. This could explain why the manipulation didn't affect the participants. Second, nominations didn't work because nominated participants did not come to perform the experiment. Thus, other methods of recruitment were used in the middle of the experiment. This contributed to the fact that participants had different academic background and were not distributed in a homogeneous way between conditions. Finally, the fact that a power analysis was not carried out may have produced a false negative since maybe there were not enough participants. Taken together, it is difficult to draw conclusions from this study. The method must be improved, specially the recruitment of participants. In the following studies, it is crucial to take into consideration the confounding variables that are affecting the results and the design should be improved to address this topic.

Chapter V

Study 3: The effect of direct and extended contact on trust in and attitudes towards social robots (part 2)

5.1. Overview

This chapter presents the second empirical study of this thesis. By taking into account all the limitations and issues encountered in the previous study, the method was improved. The main change was made to the method of recruitment. In Study 3, a power analysis was performed in order to determine the sample size. Participants had to come in pairs of friends to make sure that they knew each other. They were also allocated randomly to all conditions as they came to the robotics centre. So, unlike in Study 2, conditions were homogenous in terms of the participants' background. Apart from that, the protocol and the measures remained the same. The results of Study 3 suggest that direct contact affects both explicit and explicit attitudes but not trust. Extended contact had an effect on implicit attitudes while trust and explicit attitudes remained the same.

5.2. Introduction

This study provided one set of participants with the opportunity to interact with Pepper (Figure 11); a humanoid robot developed by Aldebaran for the Japanese Telecommunication Company SoftBank to welcome customers in their shops. Then, in order to examine the effect of direct contact on attitudes, these participants' attitudes towards robots were measured. Next, the participants were asked to record a short video describing their experience with Pepper that would be shown to a friend whom they had brought with them. This video was then shown to the participant's friend, before their attitudes towards robots were also measured; thereby providing a test of the effect of extended contact (i.e., hearing about another person's interaction) on attitudes. The findings were compared to two control conditions in which participants received neither direct nor extended contact with robots.

5.3. Method

5.3.1. Participants

Study 3 used a mixed design including time as a within-participants factor (before and after contact or extended contact), and four between-participant factors—direct contact, extended contact, no contact control, and extended contact control. A power analysis was performed to estimate the required size of the sample. We chose d = 0.48 (f^2 = 0.24) to power our study as it is just below the threshold for a medium sized effect according to Cohen's criteria. With alpha = .05 and power = 0.95, GPower 3.1 estimated the sample size needed to detect an effect of this magnitude, with a 1-within and 4-between mixed ANOVA, to be N = 80 (40 pairs, with 20 pairs per condition).

Participants were students and staff from a large University in the North of England. They were recruited using the online SONA research participation system in the Department of Psychology and via an email distribution list containing staff and students who had indicated a willingness to take part in research. Participants who were recruited via the SONA system received 4 course credits for their participation. The other participants did not receive any kind of incentive. On average, the participants were aged 23.86 (SD = 8.03); 30 of them were male and 50 were female; and the majority were British (n = 62, 78%). They came from several different backgrounds such as psychology (*N* = 29), natural sciences (*N* = 8), education (*N* = 7), healthcare (*N* = 6), engineering (*N* = 5), business (*N* = 4), computer science (*N* = 3), philosophy (*N* = 3), art (*N* = 2), fundraising (*N* = 2), languages (*N* = 2), media studies (*N* = 2), charity (*N* = 1), food industry (*N* = 1), history (*N* = 1), law (*N* = 1), physics (*N* = 1), robotics (*N* = 1) and sociology (*N* = 1).

Participants were asked to come with a friend (or someone who is closer than a friend). Once they arrived to our facilities, they were randomly allocated to one condition or another. A questionnaire was distributed in which participants had to quantify how close they felt to their friend. The scale used in this case was the Inclusion of Other in the Self scale (IOS) developed by Aron, Aron and Smollan (1992). This scale goes from 1 (to rate someone who is not close to the participant at all, like a stranger) to 7 (to rate someone who is extremely close to the participant). In addition, they were asked to write how they met. This questionnaire was kept confidential – i.e., participants were in separate rooms while completing this questionnaire and could not see what their friend had written in the questionnaire. The mean in the IOS for all the participants was 4.97 (*SD* = 1.55), which indicates that participants knew each other at least reasonably well. Out of 40 pairs of participants, 27 pairs were friends, 11 pairs were dating exclusively or married and 2 pairs were mother and daughter.

5.3.2. Measures

Explicit attitudes were measured using the Negative Attitudes towards Robots Scale (NARS) (Nomura, et al., 2006b). The NARS has 14 items (e.g., *I would feel uneasy if robots really had emotions* and *Something bad might happen if robots developed into living beings),* to which participants were asked to respond on a Likert scale from 1 to 5 where 1 corresponds to the most positive attitudes towards robots and 5 the most negative attitudes. Participants' responses were combined to create a single score reflecting their explicit attitudes towards robots (Cronbach's alpha = 0.74 and 0.64, for the measure of attitudes before and after the manipulations, respectively).

Trust in robots was measured using the human-robot trust scale, developed by Schaefer (2013). This scale presents 40 questions such as: *What percentage of the time do you think this robot will act consistently?* and *What percentage of the time do you think this robot will function successfully?* For the purposes of the present research, the questions were adapted to simply refer to "robots" (e.g., *What percentage of the time do you think robots will act consistently?*) so that they could also be answered by participants who had not met or heard about Pepper. The items were averaged to create an aggregate measure of trust where 0% represents the lowest possible level of trust and 100% represents the highest possible level of trust (Cronbach's alpha = 0.89 and 0.90, for the measure of trust before and after the manipulations, respectively).

Implicit attitudes were measured using the Implicit Association Test (IAT) (Greenwald, et al., 1998), which measures the strength of association between two target concepts (e.g., humans and robots) and two attributes (e.g., pleasant and unpleasant) by asking participants to sort words into categories as quickly as possible. The idea is that, if participants have positive attitudes towards robots, then they will be quicker to respond when the categories "robot" and "good" share a key than when "robot" and "bad" share a key. The present research used MacDorman et al.'s (2009) version of the IAT that used ten silhouettes of humans and ten silhouettes of robots as targets and eight pleasant words and eight unpleasant words as attributes. In order to obtain a value that represented participants' implicit attitudes, a D score was calculated using the algorithm described by Greenwald et al. (2003); higher D scores reflect more negative implicit attitudes towards robots.

5.3.3. Robot used

The aim of this study is to investigate attitudes towards social humanoid robots. Therefore, it was necessary that participants interacted with a social humanoid robot. As in the previous study, Pepper was the robot used (Figure 11). This robot is user friendly and has all the necessary capabilities to perform in an experiment like this. In fact, it is capable to have a conversation and respond appropriately according to the participant's reactions. The programming that was implemented was the same as in the previous study.

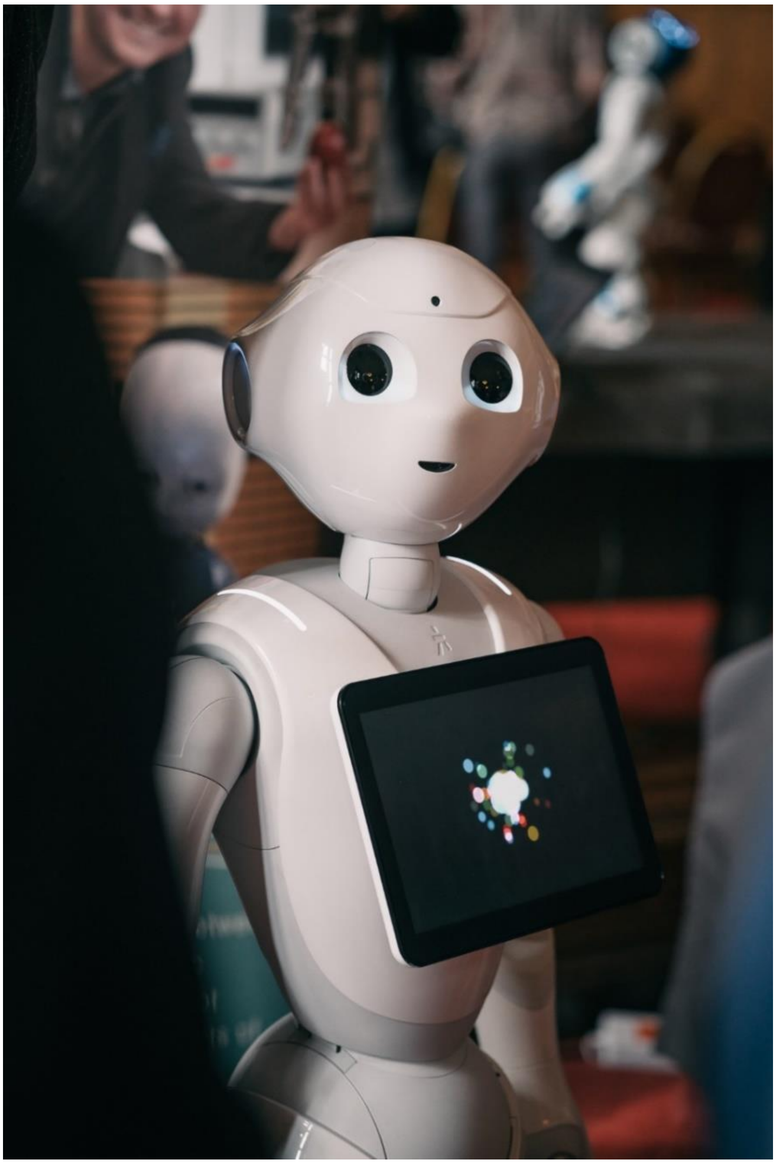


Figure 11: The humanoid robot Pepper. Height: 120cm.

5.3.4. Procedure

Ethical approval was obtained by Department of Psychology Research Ethics Committee. Prior to taking part in the main study, all of the participants completed an online questionnaire that asked for their demographic information and measured their explicit attitudes towards and trust in robots before signing up to attend an experimental session at a centre for robotics research with someone that they considered a friend or closer. On arrival, participants completed the IOS and a measure of implicit attitudes towards robots before being randomly allocated to an experimental or control condition.

There were two **experimental conditions**. Participants in the **direct contact (DC) condition** interacted directly with the robot. Participants watched an instructional video (https://youtu.be/-7FjY8XE5N8) that asked them to pretend that they were in a shop that contained a range of hair care products and that Pepper was the shop assistant. The conversation was limited to five minutes. After the interaction, the participants completed the measures of trust, explicit and implicit attitudes again. Finally, they recorded a short video message for their friend describing their experience with the robot. Specifically, participants were asked to answer the following questions while talking to the camera:

What happened since you first saw Pepper until the end of the interaction?

What did you talk about?

How did you feel while interacting with Pepper?

Did Pepper help you to achieve the purpose of the conversation (i.e., choosing the product)?

Did you like Pepper? Why?

Participants then left the room and their friend, who was in the **extended contact (EC) condition**, came in. Participants in the EC condition watched the video that their friend had recorded and then completed the trust, explicit and implicit measures of attitudes a second time.

At the same time, there were two **control conditions**. Participants in the **control (C) condition** did not interact with the robot or watch a video. It is also worth remembering that these participants only completed the IAT once like they did in the previous study. These participants simply completed the measures of trust, explicit and implicit attitudes. In order to control for the effects of watching a video made by a friend, participants in the control condition were instructed to record a short video once they had finished with all the questionnaires and the IAT. In this video, they were talking about someone that they had met recently for the first time (e.g., a new friend, a shop assistant, a waiter, a taxi driver, a receptionist). Specifically, participants were asked to answer the following questions while talking to the camera:

> *What happened with this person since you began this conversation until you finished talking?*
>
> *What did you talk about?*
>
> *How did you feel interacting with this person?*
>
> *If the conversation had a purpose did this person help you to achieve the purpose of the conversation?*
>
> *Did you like this person? Why?*

Because participants in the C condition recorded this video once all their data was collected, this did not affect their answers in the questionnaires or the IAT. Participants then left the room and their friend, who was in the **extended contact control (ECC) condition**, came in. Participants in the ECC condition watched the video that their friend had recorded talking about someone they had met recently for the first time and then completed the measures of explicit and implicit attitudes a second time.

All participants received a debrief form after performing the experiment and they also had the change to ask any question to the researcher.

5.4. Results

Skewness and kurtosis were used in order to test the normality of the data. All the values fell between -1.96 and +1.96, which is considered acceptable in order to prove normal univariate distribution (George, 2011). This is shown in Table 10, which also shows the effect of condition and time on each of the measures of attitude (see also Figures 12, 13, and 14).

5.4.1. Randomization check and correlations between measures.

A series of one way independent ANOVAs were conducted in order to check that the attitudes of participants in the four conditions did not differ at baseline. Consistent with the idea that randomization was successful, there were no significant differences in explicit attitudes, $F(3, 76) = 0.10$, $p = .960$, partial eta^2 = .00, trust, $F(3, 76) = 0.27$, $p = .846$, partial eta^2 =.01, or implicit attitudes, $F(3, 76) = 0.63$, $p = .598$, partial eta^2 = .02, between the conditions at baseline. There was a negative correlation between the NARS and the Trust scale ($r = -0.47$, $n = 160$, $p < 0.001$) which means that the more negative attitudes towards robots, the less trust people have in robots. There was no correlation between the IAT and any other measure.

5.4.2. Effect of direct and indirect contact on attitudes

In order to examine the effects of direct and indirect contact (relative to control procedures) on attitudes towards robots and trust, we conducted two 4-between (condition: direct contact, extended contact, extended contact control, control) by 2-within (time: before vs. after) mixed ANOVAs with explicit attitudes and trust as the dependent variables. This was done primarily to find an interaction between condition and time although main effects are also reported.

The main effect of condition on explicit attitudes towards robots was not significant, $F(3, 76) = 0.27$, $p = .850$, partial $\text{eta}^2 = .01$. However, there was a significant effect of time on explicit attitudes, $F(1, 76) = 7.15$, $p = .009$, partial $\text{eta}^2 = .09$, that was qualified by a significant interaction between condition and time, $F(3, 76) = 3.56$, $p = .018$, partial $\text{eta}^2 = .12$. A series of paired samples t-tests indicated that explicit attitudes changed as a function of direct contact, $t(19) = 2.86$, $p = .010$; such that participants had more positive explicit attitudes towards robots after interacting with Pepper. Explicit attitudes did not change significantly as a function of extended contact, $t(19) = 0.24$, $p = .817$, or either of the control procedures: ECC condition, $t(19) = 1.07$, $p = .300$, and C condition, $t(19) = 0.06$, $p = .949$.

The same analysis with trust as the dependent variable revealed a main effect of time, $F(1, 76) = 5.23$, $p = .025$, partial $\text{eta}^2 = .06$. Participants had more trust in robots after taking part in the study ($M = 61.02$, $SD = 11.40$) than before ($M = 58.67$, $SD = 12.34$). However, the effect of condition, $F(3, 76) = 0.41$, $p = .749$, partial $\text{eta}^2 = .02$, and the interaction between condition and time, $F(3, 76) = 0.09$, $p = .966$, partial $\text{eta}^2 = .00$, were not significant.

Finally, we conducted a one-way ANOVA to examine the effect of condition on implicit attitudes following the procedures.[2] There was a significant main effect of condition on implicit attitudes, $F(3, 76) = 3.00$, $p = .036$, partial eta$^2 = .11$; participants in the extended contact condition had more positive implicit attitudes than those in the extended contact control condition. We also ran a series of paired samples t-test to examine the effect of time on implicit attitudes in the conditions that completed the IAT before and after contact or control procedures. These analyses suggested that implicit attitudes became more positive as a result of direct contact, $t(19) = 3.05$, $p = .007$, and extended contact, $t(19) = 2.49$, $p = .022$, but not as a result of extended contact control procedures, $t(19) = 0.93$, $p = .364$.

[2] It was not possible to run a repeated measures ANOVA as implicit attitudes were only measured once in one of the control conditions.

Table 10

Study 3: Descriptive statistics. Effect of condition and time on each of the measures of attitude.

		Direct contact						Extended contact					
		Explicit attitudes		Trust		Implicit attitudes		Explicit attitudes		Trust		Implicit attitudes	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Mean	2.82*	2.48*	56.55	58.41	0.50*	0.31*	2.75	2.74	59.17	61.79	0.41*	0.25*
	SD	0.70	0.51	11.42	9.95	0.35	0.38	0.66	0.63	14.99	11.74	0.36	0.32
Skewness	Statistic	0.54	-0.01	-0.82	0.12	-0.54	0.32	0.45	-0.26	-0.17	-0.46	-0.13	0.98
	Std. Error	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
Kurtosis	Statistic	1.49	-0.50	0.49	-0.87	0.63	-0.47	-0.18	-0.07	0.64	0.17	-0.31	1.51
	Std. Error	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99

Note. * indicates a significant ($p < .05$) difference between before and after assessments.

Table 10 (continuation)

Study 3: Descriptive statistics. Effect of condition and time on each of the measures of attitude.

		Extended contact control						Control					
		Explicit attitudes		Trust		Implicit attitudes		Explicit attitudes		Trust		Implicit attitudes	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Mean	2.72	2.64	59.88	61.71	0.44	0.34	2.80	2.79	59.08	62.18	0.54	-
	SD	0.53	0.62	11.71	11.41	0.36	0.30	0.53	0.57	11.59	12.78	0.30	-
Skewness	Statistic	0.39	0.30	-0.67	0.18	-0.48	0.19	1.28	0.74	0.17	0.00	0.06	-
	Std. Error	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	-
Kurtosis	Statistic	-0.79	-0.35	0.58	-1.07	-0.07	-0.52	0.97	0.25	-0.25	-0.32	0.16	-
	Std. Error	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	-

Note. * indicates a significant ($p < .05$) difference between before and after assessments.

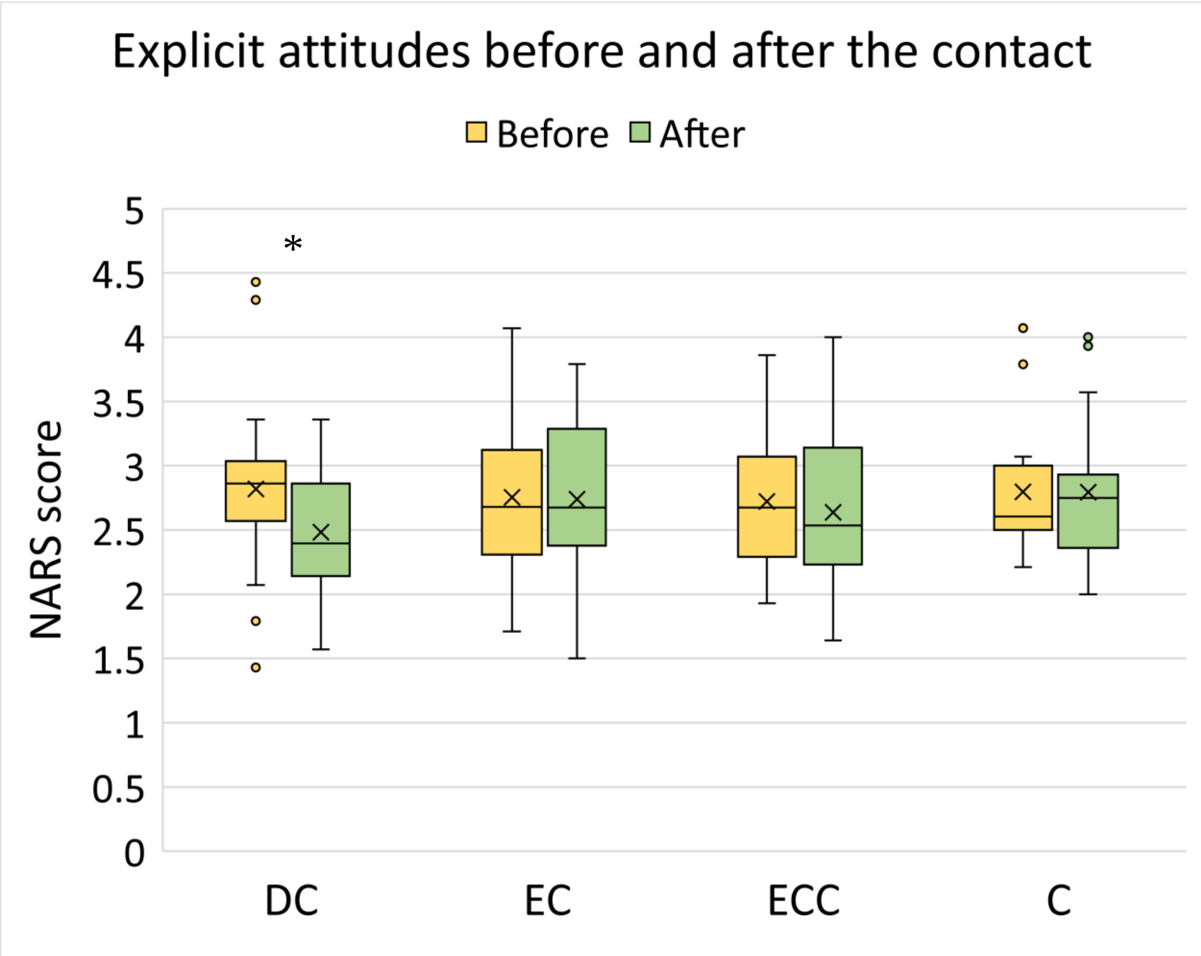


Figure 12: Effect of Condition (DC = Direct Contact, EC = Extended Contact, ECC = Extended Contact Control, C = Control) on Explicit Attitudes towards Robots (NARS) (Study 3). * p < 0.05

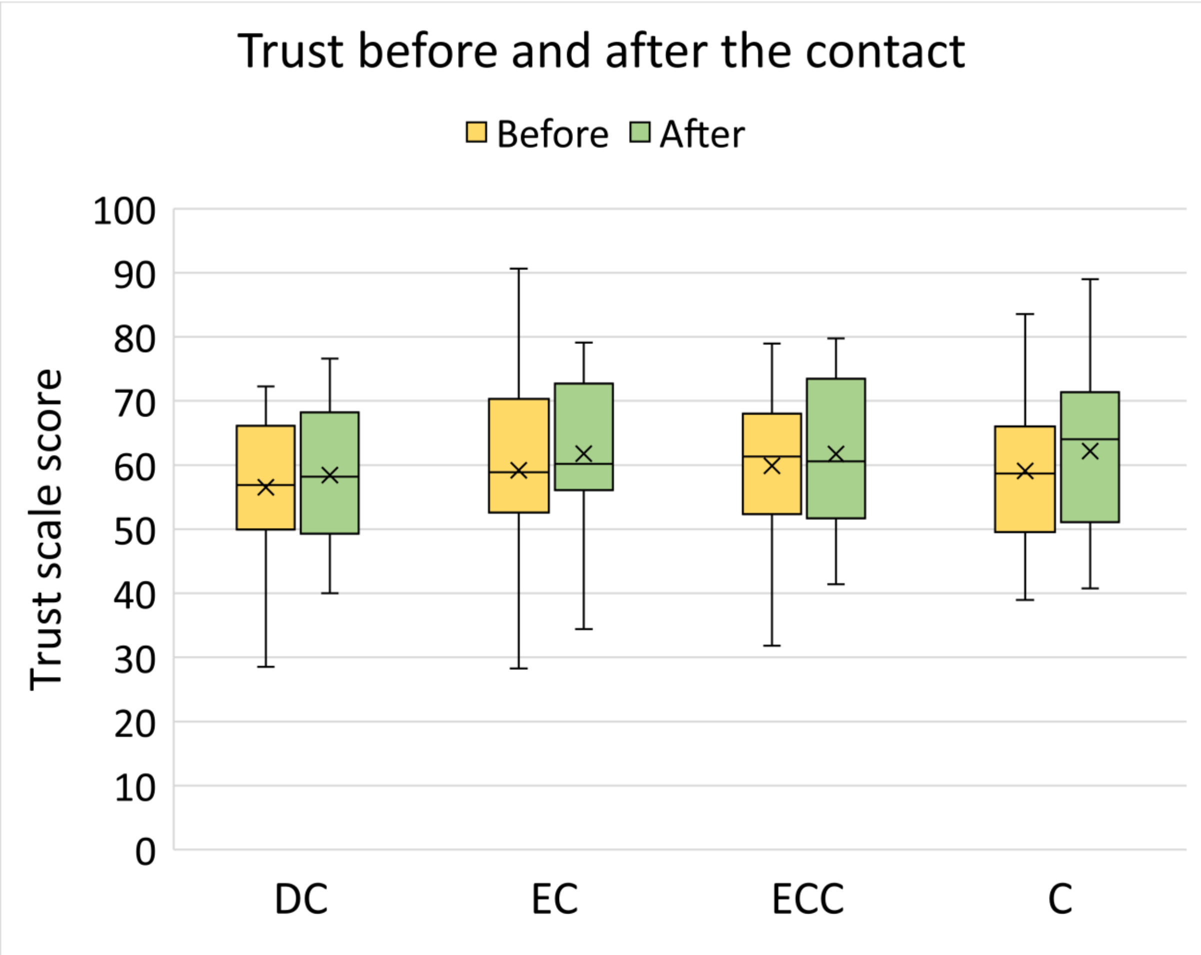


Figure 13: Effect of Condition (DC = Direct Contact, EC = Extended Contact, ECC = Extended Contact Control, C = Control) on Trust in Robots (Study 3). * $p < 0.05$

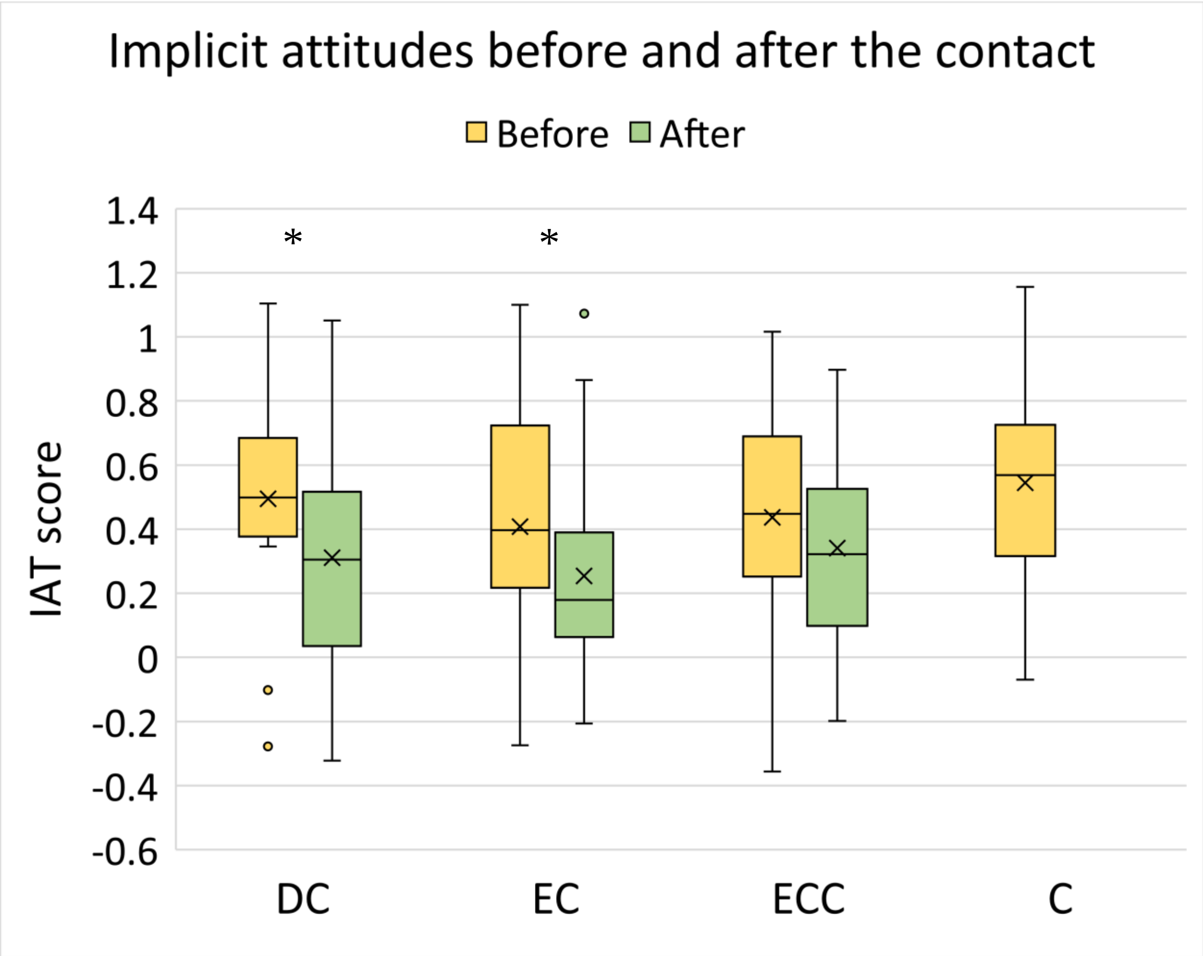


Figure 14: Effect of Condition (DC = Direct Contact, EC = Extended Contact, ECC = Extended Contact Control, C = Control) on Implicit Attitudes (Study 3). * $p < 0.05$

5.5. Discussion

The findings of Study 3 suggested that direct contact with a social humanoid robot changed participants' explicit and implicit attitudes towards robots, but did not affect levels of trust. Extended contact changed implicit attitudes (i.e., participants evidenced more positive implicit attitudes after watching a video of a friend talking about their experience with Pepper), but there was no evidence that extended contact influenced explicit attitudes or trust. Taken together these findings provide support for extending the contact hypothesis (Allport, 1954) to human-robot interaction. However, we did not predict that extended contact would influence implicit, but not explicit attitudes. We therefore decided to carry out a (pre-registered) conceptual replication of the extended contact procedures developed in Study 3 to see if we obtained the same findings before drawing conclusions.

Chapter VI

Study 4: The effect of direct and extended contact on trust in and attitudes towards social robots (conceptual replication)

6.1. Overview

This is the last empirical research of this thesis. It is a conceptual replication of the previous study. Some of the results obtained in Study 3 were unexpected. Participants in the EC condition changed their implicit attitudes after the intervention but they did not change their explicit attitudes. In order to validate this results, we decided to carry out a conceptual replication by only examining extended contact in contrast to a control condition with no contact. The results were replicated; extended contact had an effect on implicit attitudes but not explicit attitudes or trust. The discussion section of this chapter explores the implications of these findings and provides several explanations that could justify them.

6.2. Introduction

Study 4 replicated the procedures of Study 3, except that the humanoid robot NAO (depicted in Figure 15) developed by Aldebaran Robotics in 2008, was used in place of Pepper and we only examined the effects of extended contact, relative to no contact. NAO was presented in a cinema setting in which the robot recommended films to participants. These changes were made because the results should not have been linked to a specific robot or a specific task. The robot and the task were changed with the intention to eliminate the possibility of having extra confounding variables that could affect the results.

The procedures and approach to analysis were pre-registered on AsPredicted.org (#17464) before starting the data collection. During the preregistration, it was necessary to provide an exhaustive explanation of all the details of the experimental protocol and provide information about the analysis they want to carry out after collecting the data.

6.3. Method

6.3.1. Participants

Study 4 used a mixed design including time as a within-participants factor (before and after extended contact), and two between-participant factors—no contact and extended contact. A power analysis was performed to estimate the required size of the sample based on a medium-sized difference between extended contact and control conditions (d = 0.48, which equates to effect size f^2 = 0.24) as we did in the previous study, but with a lower power threshold (0.80) since study 3 gave us more confidence of an effect. With alpha = .05, power = 0.80, and two conditions, GPower 3.1 recommended a sample size of N = 38, or 19 pairs. As in the previous study, participants were students and staff from a large university in the North of England. They were recruited using the online SONA research participation system and via an email distribution list containing staff and students who had indicated a willingness to take part in research. Participants who were recruited via the SONA system received 4 course credits. The other participants did not receive any compensation. Participants came in pairs; 46 participants took part in the study (23 pairs). Their mean age was 23.24 (SD = 10.14); 15 of them were male and 31 were female; and the majority were British (N = 36, 78%). They came from several academic or professional backgrounds such as psychology ($N = 20$), business ($N = 5$), natural sciences ($N = 4$), computer science ($N = 3$), healthcare ($N = 3$), anthropology ($N = 2$), languages ($N = 2$), art ($N = 1$), education ($N = 1$), law ($N = 1$), religion ($N = 1$), robotics ($N = 1$), sports ($N = 1$) and one of them had just graduated high school ($N = 1$).

Like in Study 3, participants were asked to come in pairs of friends (or closer than friends). Once they arrived to our facilities, they were randomly allocated to one condition or another. They were asked to fill the IOS (Aron et al., 1992), in which they had to quantify how close they felt to their friend, following the same procedure described in Study 3. The mean in the IOS for all the participants was 4.89 ($SD = 1.50$), which indicates that participants knew each other. Out of the 23 pairs of participants, 18 pairs were friends, 4 pairs were dating exclusively or married and 1 pair were mother and daughter.

6.3.2. Measures

Since Study 4 was a replication of Study 3, the same measures were used in order to test participants' attitudes towards robots. These are the Negative Attitudes towards Robots Scale (NARS) (Nomura, et al., 2006b), the Trust scale by Schaefer (2013) and the Implicit Association Test (IAT) (Greenwald, et al., 1998; MacDorman et al., 2009).

6.3.3. Robot used

Like in Study 3, the aim was to investigate attitudes towards social humanoid robots. Nonetheless, it was important that the results were not tied to a specific robot. Therefore, it was better to use another robot other than Pepper. This implied that there was a need to have another robot that was also a social humanoid robot. For this reason, NAO (Figure 15) was the chosen robot to perform this experiment. This robot is a humanoid that was also designed to interact with people. Apart from that, NAO has the same user interface as Pepper, which made the programming of the robot more convenient. The code was adapted from the previous studies and this time, the robot was programmed to discuss films. For a full list of keywords see Supplementary Material 4.

Figure 15: The humanoid robot NAO. Height: 57cm.

6.3.4. Procedure

Ethical approval was obtained by Department of Psychology Research Ethics Committee. As in the previous study, participants completed online measures of their explicit attitudes towards and trust in robots (Cronbach's alpha NARS = 0.70, Trust = 0.88) before registering to take part in the main study alongside someone that they considered a friend or closer. On arrival, the participants completed the IOS. Then, they were randomly allocated to either the control condition or the extended contact condition and asked to wait in separate rooms.

This time, the study only had 2 conditions in contrast to the previous study. This was done because the analysis was only done to the effect of extended contact on attitudes towards robots. That is to say, there was not DC condition because direct contact has already been proved empirically to have an effect on attitudes towards robots. For this reason, it was not the focus of this study. Moreover, there was not ECC condition. Instead, there was a control condition with no interaction with robots and no video. This decision was made because, in the previous study, there were no significant differences between the ECC condition and the control condition.

Participants who were allocated to the **control (C) condition** were asked to complete the same Implicit Association Test (IAT) as used in Study 3, along with the same measures of their explicit attitudes and trust in robots (Cronbach's alpha NARS = 0.68, Trust = 0.90). Again, all participants used the same computer in the laboratory to complete the IAT. No one could have done the IAT before coming to perform the experiment. This means that participants in the C condition only completed the IAT once because, taking into account that they did not interact with anything, it would not have made sense for them to complete the IAT twice in a row.

Once these participants had finished completing the questionnaires and tasks, all the data for this condition had been collected. Nonetheless, participants in the C condition were then asked to interact with NAO so that they could record a video describing their experiences which would allow us to examine the effect of extended contact by showing it to their friend. This means that no data was collected after the interaction with the robot because, as said previously, direct contact was not analysed in this study. As before, participants watched an instructional video on how to interact with the robot (https://youtu.be/AexluOIholc) and then interacted with NAO for 5 minutes; during which time NAO asked questions about the participant's taste in films and made some recommendations.

When the interaction was finished, the participant was given similar instructions as in Study 3 to record a video talking about their interaction with NAO. After each participant had recorded their video, they left the room and their friend, who was in the Extended Contact (EC) condition, entered the room.

Participants in the **extended contact (EC) condition** completed the IAT and then watched the video that their friend had recorded previously, in which they described their interaction with the robot. This time, the person in the video had interacted with NAO instead of Pepper. Apart from that, the interaction had been different; instead of recommending hair care products, the robot was recommending films. These changes were made because the results needed to be detached from a specific robot or a specific interaction with the robot. After watching the video, participants in the EC condition completed the IAT again and the questionnaires measuring their explicit attitudes and trust in robots. Finally, both participants were debriefed and the researcher answered any questions that they had.

6.4. Results

Table 11 and Figures 16, 17, and 18 show the effect of condition and time on each of the measures of attitude. Two mixed ANOVAs were used to determine if explicit attitudes and trust were affected by condition and time (and their interaction).

There were no significant effects of condition, $F(1, 44) = 1.17$, $p = .285$, partial eta^2 = .03, time, $F(1, 44) = 1.24$, $p = .272$, partial eta^2 = .03, or the interaction between condition and time, $F(1, 44) = 0.29$, $p = .592$, partial eta^2 = .01, on explicit attitudes as measured by the NARS.

There was no significant effect of condition, $F(1, 44) = 0.90$, $p = .346$, partial eta^2 = .02, or an interaction between condition and time, $F(1, 44) = 0.20$, $p = .658$, partial eta^2 = .01, on trust, but there was a significant effect of time, $F(1, 44) = 8.44$, $p = .006$, partial eta^2 = .16. Participants had more trust in robots after taking part in the study ($M = 58.94$, $SD = 11.81$) than before ($M = 62.32$, SD =12.21).

We also conducted two t-tests to examine the effect of condition on implicit attitudes at baseline (where we assumed that there would be no differences) and then the effect of time in the EC condition, which reflects the effect of extended contact on implicit attitudes. There was no difference in implicit attitudes between the conditions at baseline, $t(44) = -0.10$, $p = .921$; however, as expected, there was a significant effect of time on IAT scores in the extended contact condition, $t(22) = 2.45$, $p = .023$, suggesting that extended contact led participants to hold more positive implicit attitudes.

As in the previous study, there was a negative correlation between the NARS and the Trust scale ($r = -0.26$, $n = 92$, $p = 0.01$) which means that the more negative attitudes towards robots, the less trust people have in robots. There was no correlation between the IAT and any other measure.

Table 11

Study 4: Descriptive statistics

		Control						Extended contact					
		Explicit attitudes		Trust		Implicit attitudes		Explicit attitudes		Trust		Implicit attitudes	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
	Mean	2.65	2.57	60.80	63.66	0.55	-	2.79	2.76	57.09	60.98	0.56*	0.38*
	SD	0.59	0.60	10.74	11.71	0.23	-	0.48	0.42	12.76	12.82	0.25	0.34
Skewness	Statistic	-0.74	-0.60	-0.06	0.12	-0.08	-	0.49	0.69	0.61	-0.32	-0.04	-0.44
	Std. Error	0.48	0.48	0.48	0.48	0.48	-	0.48	0.48	0.48	0.48	0.48	0.48
Kurtosis	Statistic	0.53	-0.11	0.33	0.44	-0.96	-	-0.86	0.14	-0.02	-0.39	-1.00	-0.37
	Std. Error	0.93	0.93	0.93	0.93	0.93	-	0.93	0.93	0.93	0.93	0.93	0.93

Note. * indicates a significant ($p < .05$) difference between before and after assessments.

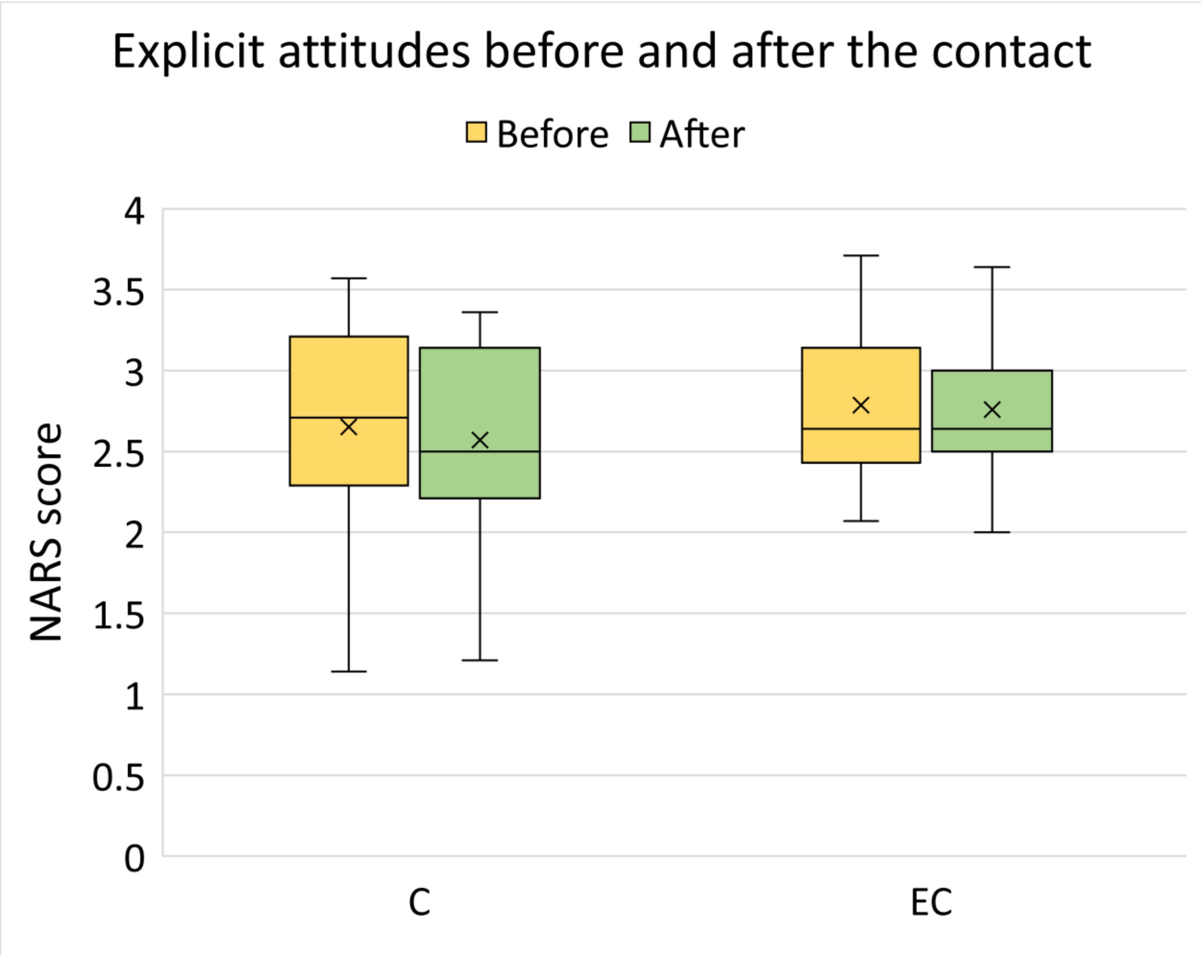


Figure 16: Effect of Condition (C = Control, EC = Extended Contact) on Explicit Attitudes towards Robots (NARS) (Study 4). * $p < 0.05$

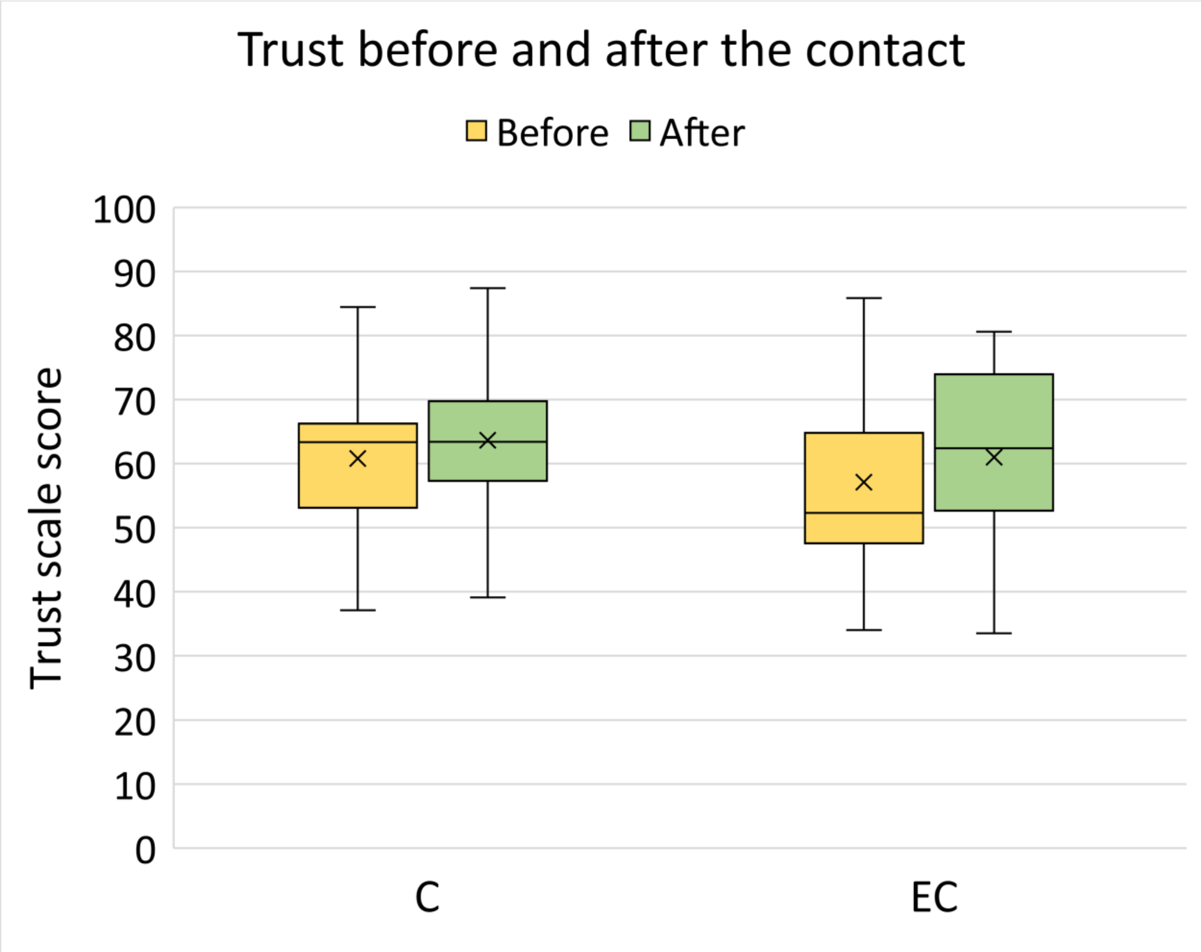


Figure 17: Effect of Condition (C = Control, EC = Extended Contact) on Trust in Robots (Study 4). * $p < 0.05$

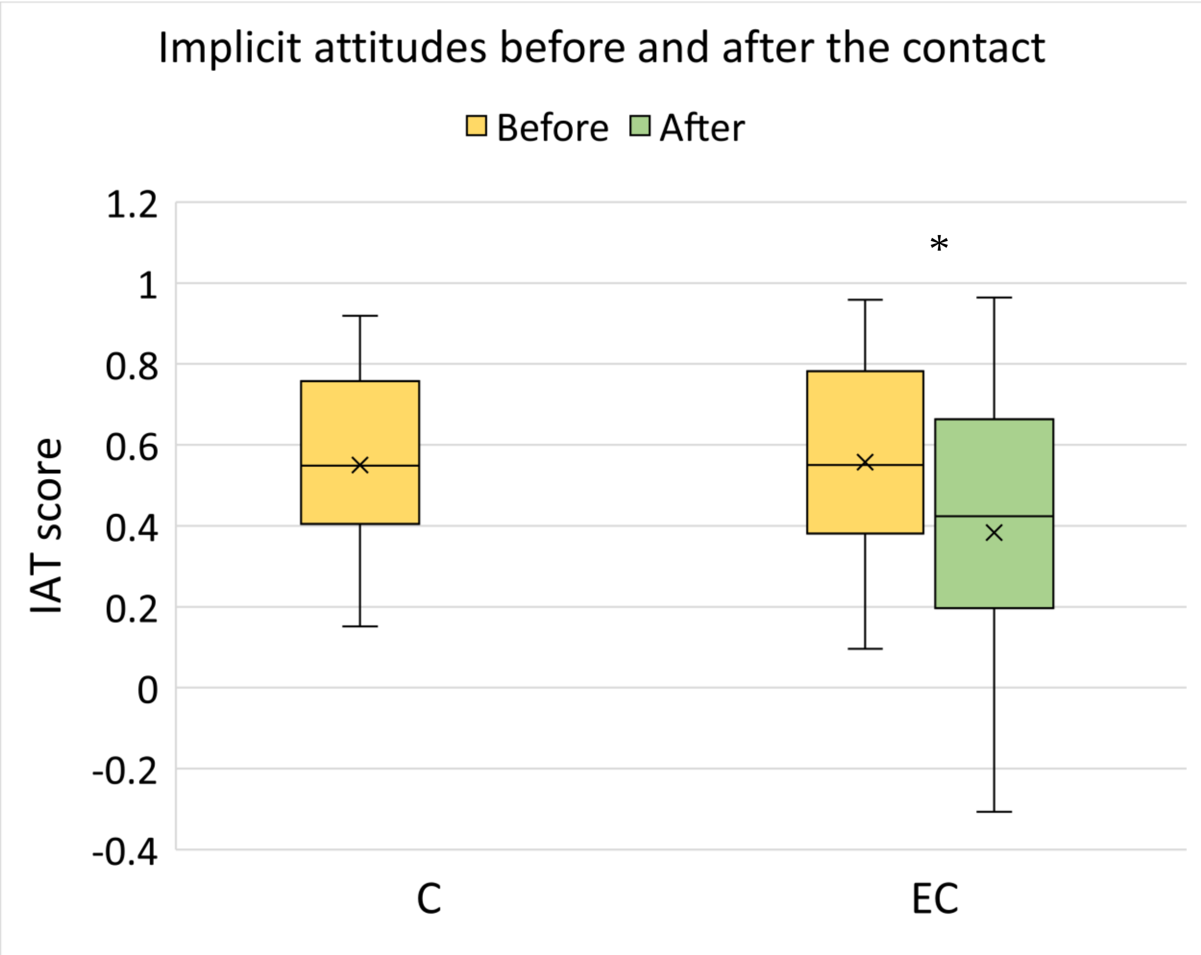


Figure 18: Effect of Condition (C = Control, EC = Extended Contact) on Implicit Attitudes towards Robots (Study 4). * $p < 0.05$

6.5. Discussion

The present research drew on the psychology of intergroup relations as the inspiration for investigating how people respond to new 'social' technologies; in this case, social robots. Two studies found evidence that direct contact with a social humanoid robot had a positive effect on participants' (explicit and implicit) attitudes towards robots. Taken together with previous studies showing positive effects of direct contact with robots on explicit attitudes (e.g., Nomura et al., 2008; 2011), these findings suggest that Allport's (1954) contact hypothesis can be extended to non-human agents and point to the potential of using research on how groups of humans interact with and think about one another, to understand how humans interact with and think about artificial social agents.

The results obtained both in the main study and its replication suggested that having extended contact with a humanoid robot (i.e., watching a friend explaining their interaction with the robot) positively changed implicit attitudes while explicit attitudes remained the same. In contrast, having direct contact with a humanoid robot changed both explicit and implicit attitudes, making them more positive. These results were unexpected but several explanations could explain this discrepancy between implicit and explicit measures. Previous research suggests that disagreement between explicit and implicit measures is not uncommon. Some studies revealed that explicit attitudes can be changed more easily than implicit attitudes (Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Brinol, & Jarvis, 2006), while others show the opposite results (Barden, Maddux, Petty, & Brewer, 2004; Dasgupta & Greenwald, 2001; Wittenbrink, Judd, & Park, 2001). Although it seems that there is no consistency among these studies, they all agree that implicit and explicit measures do not necessarily correlate. The main plausible explanations for this incongruity are discussed in the following sections.

6.5.1. Subliminal priming

One reason that could explain a change in implicit attitudes but not in explicit attitudes is that a stimulus is presented outside conscious awareness and therefore, it only affects participants unconsciously (Rydell, McConnell, Mackie, & Strain, 2006). Unconscious stimuli (also known as subliminal stimuli) are, for example, images that are flashed so quickly that the viewer cannot perceive consciously. Nonetheless, it is essential to understand that subliminal stimuli are, by definition, imperceptible by the conscious mind and, therefore, they are impossible to find by the people who are exposed to them. That is to say, people are not aware of the stimuli they are receiving. The stimulus used in the empirical research in this thesis was extended contact presented in a form of a video in which participants watched their friend talking about their previous interaction with a robot. This kind of stimuli was not outside conscious awareness; in fact, it is fairly explicit. So, participants should have been consciously aware of what they were watching. This means that participants could not have been affected by subliminal priming.

6.5.2. Supraliminal priming

Although participants were aware of the stimuli that they were exposed to (i.e., the video of their friend talking about the robot) in the studies reported in this thesis, they may not have been aware of the way in which it was influencing their attitudes. That is to say, the video used to generate extended contact, may have acted as a form of supraliminal priming that affected the participants’ implicit, but not explicit, attitudes.

> In supraliminal priming, the participant is exposed to the priming stimuli as part of a conscious task. That is, the individual is fully aware of the priming stimuli itself but is not aware of some underlying pattern that serves to prime to prime the construct (Bargh & Chartrand, 2000, pp. 317).

An example of supraliminal stimuli is the background music that people hear in shops or supermarkets. Customers can perceive consciously that there is music in the background. However, they are not typically aware that it is influencing their buying behaviour. For example, Milliman (1982) investigated how different tempos in the background music affected sales in a supermarket. Their findings indicated that slow tempo music increased sales by 38.2% compared to fast tempo music. In addition, the music also affected the pace of in-store traffic; customers moved slower around the supermarket when they were hearing slow tempo music, thing that could have been influencing this increase on the sales.

To give another example of supraliminal priming, some words can be used to elicit a reaction or maybe modify the behaviour of a group of participants. Earlier priming investigations used this method to analyse if priming a trait category could affect the way participants would perceive a certain individual (Bargh & Pietromonaco, 1982; DeCoster & Claypool, 2004; Fazio, 2001; Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979). For example, Higgins et al. (1977) asked participants to do two different tasks that were allegedly unrelated. The first task consisted of memorizing several words with either positive or negative connotations. In the second task, participants were asked to read about a fictional character called Donald, who was portrayed as a person who liked to do exciting activities. In this way, participants could form an opinion about him. Then, they were asked to answer a questionnaire in which they evaluated Donald. Results show that participants who, at the beginning of the experiment, were exposed to positive words rated Donald more positively than participants who were exposed to negative words.

Srull and Wyer (1979) performed a similar study using a priming technique called the Scrambled Sentence Test (SST), developed originally by Costin (1969). This time, participants did not have to memorize words but were presented with some scrambled words and had to be arranged in the correct grammatical order to form a sentence. After that, they had to read a text about Donald and answer some questionnaires like in the study mentioned previously. The findings supported the research done previously and they extracted the following conclusions:

> [O]nce a trait concept or schema is made more accessible by previous cognitive activity, the likelihood that the same schema will be used to encode new information is increased. The accessibility of these concepts, and therefore the likelihood that they are subsequently used, increases with the number of times that instances of them have been activated in the past. […] [O]nce behavioral information is encoded, these encodings affect judgments of the person who manifested the behavior with respect to both the trait originally primed and other traits that are related to it only indirectly through subjects' implicit personality theories. (Srull & Wyer, 1979, pp. 1669-1670)

In other words, participants' social judgement of another individual was affected by being previously exposed to several words which had certain connotations.

These previous studies provide examples of supraliminal priming, but they only analyse its effect on explicit self-reported measures (e.g., questionnaires) or behavioural measures (e.g. spending more money in a supermarket). Nonetheless, in the studies reported in the present thesis, participants who were exposed to extended contact with the robot reported a change in implicit attitudes but not explicit attitudes. Unfortunately, little research has investigated how supraliminal priming affects both explicit and implicit attitudes. Skinner and Cheadle (2016) investigated the “Obama effect” (i.e. how Obama winning the elections in the USA affected racism) and used supraliminal priming with the intention to modify implicit racial bias. They had two experimental conditions in which they used two kinds of stimuli in order to prime either inter-group power threat or inter-group majority threat, and a control condition with no priming. In order to manipulate power threat, participants read an article about the historic importance of Obama in the elections. Researchers hypothesised that white participants would feel threatened by the fact that the president of the USA is representative of a minority group. Similarly, they manipulated majority threat by asking participants read another article talking about the minority-majority population shift in the USA. Again, they hypothesised that white participants would feel threatened by this fact and, thus, present racial bias in the IAT. Only participants who already presented racial bias before performing the experiment were affected by the power threat, magnifying their bias in the IAT after reading the article. In contrast, all participants were influenced by the majority threat, showing an increased racial bias in the IAT after regain the article. These results suggest that not only the IAT is affected not only by supraliminal priming but also participants’ preconceptions and prejudices.

Penn (2016) investigated the effects of advertising on the explicit and implicit attitudes regarding a specific company, in this case; eBay. Participants were divided into regular users of eBay and non-users. One half of the participants were exposed to a video advertising the platform, while the other half (the control condition) were not exposed to any stimulus. After that, researchers assessed participants’ attitudes towards eBay using both explicit and implicit measures. They reported that participants who were non-users of eBay and were exposed to the advertising show significantly more positive implicit attitudes after the intervention. In this way, the video supraliminally primed participants' implicit attitudes towards eBay.

Supraliminal priming can actually be related to the stimuli that were presented to the participants in the empirical studies reported in this thesis. Specifically, participants were aware of the video that they were watching but may not have been aware of how this could affect their attitudes. In a way, watching the video of a friend talking about the robot might have had a priming effect since the videos had positive words about the robot and, therefore, it may have affected participants' implicit attitudes. In the same way that participants, in one of the studies mentioned previously (Higgins et al., 1977), rated Donald more positively after being exposed to positive words, participants watching a video of a friend who talks positively about the robot could have been affected by this priming effect and changed their implicit attitudes towards robots in a more favourable way. In fact, previous studies support the idea that implicit attitudes can be modified by recent experiences and that this can happen with no conscious control (Dasgupta & Greenwald, 2001; Lowery, Hardin, & Sinclair, 2001; Mitchell, Nosek, & Banaji, 2003).

6.5.3. Defensive reactions

Another possible explanation for the discrepancy between implicit and explicit attitudes in the studies reported in this thesis could be that the IAT can measure what people can't or won't say directly. Maybe participants interpreted extended contact as intended to persuade them and so resisted to change their opinion in the questionnaire in an attempt to be consistent with their previous answers. There is evidence suggesting that extended contact can be potentially persuasive (LaCour & Green, 2014). In addition, evidence also suggests that people are reluctant to change their opinion when they are being persuaded with evidence and especially once they have already expressed their point of view (Hart et al., 2009), for example, in a questionnaire. Zajonc made the following statement: “behaviour and attitudes are not only consistent to the objective observer, but that individuals try to appear consistent to themselves.” (Zajonc, 2017, pp. 63) According to dissonance theory, factors or stimuli that cause cognitive dissonance elicit defence motivations in order for the individual to remain consistent (Hart et al., 2009). These mechanisms are strongly related to the bias of congeniality (also known as confirmation bias), which is the tendency to interpret or select new information that confirms our previous beliefs (Plous, 1993). This defence motivation would be strengthened when people who had already expressed their attitudes receive challenging information related to the topic in question (Frey, 1986). That is, when people are confronted by information which challenges their attitudes, beliefs, or behaviours that have recently been expressed, their effort to reduce cognitive conflict can reinforce the bias of congeniality. Taken together, because participants had already positioned themselves by answering the questionnaire before the intervention, they may have felt some degree of discomfort if they had changed their opinion in a matter of minutes. However, because implicit attitudes are not controlled by the conscious mind (Greenwald & Banaji, 1995), the IAT detected the change in attitudes that could not possibly be reported explicitly since that would create cognitive dissonance.

6.5.4. Implicit attitudes are more fluctuant than explicit attitudes

A series of studies carried out by Gawronski, Morrison, Phills and Galdi (2017) suggested that implicit attitudes are more fluctuant and sensitive to recent experiences than explicit attitudes. In this research, implicit measures showed lower levels of temporal stability than explicit measures. Apart from that, implicit measures also showed less resistance to change by the influence of recent experiences. Therefore, it could be the case that, in the present research, participants who were exposed to extended contact had their implicit attitudes affected momentarily. Nonetheless, up until now, there is no other research validating or replicating these results. Therefore, it would be wise to take the conclusions obtained by Gawronski et al. (2017) with caution.

6.5.5. Explicit and implicit attitudes are not mediated by each other

Lastly, it could also be possible that the correlation between implicit and explicit attitudes varies depending on the task or the topic and, although researchers usually observe a correlation, explicit and explicit attitudes may not be mediated by each other. Evidence suggests that implicit attitudes can be modified by brief experiences without having an effect on explicit attitudes (Blair, Ma & Lenton, 2001; Dasgupta & Greenwald, 2001, Kawakami, Dovidio, Moll, Hermsen & Russin, 2000). Indeed, a meta-analysis of the effect of different methods changing implicit attitudes by Forscher et al. (2019) “found little evidence that changes in implicit measures translated into changes in explicit measures” (Forscher et al., 2019, p. 544-545). This would explain why, sometimes, there is no relationship between implicit and explicit attitudes. It could be the case that they are not always necessarily related.

6.6. Limitations and Future Directions

The studies described here have a number of strengths – they are theoretically motivated, adopt experimental designs, use established measures of attitudes, and demonstrate consistent and robust effects. Furthermore, the conceptual replication of the effects of extended contact on implicit attitudes using different robots suggests that the findings may be extrapolated to direct and indirect contact with other kinds of social robots. One limitation of the studies, however, is that attitudes were measured immediately following the contact procedures and we did not consider whether participants' attitudes shaped their subsequent behaviour. Therefore, it is unknown if the changes in attitudes that were observed as a function of contact last over time or influence behaviour. Similar procedures have been shown to promote relatively enduring changes in participants' attitudes towards groups of humans (e.g., Eller & Abrams, 2004), as well as promote positive expectations about interactions and responses during actual interactions (e.g., Mallett & Wilson, 2010; West & Turner, 2014), suggesting that direct and indirect contact with social robots may produce lasting and meaningful changes in attitudes; however, longitudinal studies need to confirm this.

Another limitation in the replication study is that the control condition that was used had no intervention at all. This decision was taken based on the fact that, in the previous study, there were no significant differences between control conditions. However, this issue needs to be mentioned here because a non-relevant EC condition could have controlled for other potential explanations for observed differences.

6.7. Conclusion

The present research directly tackles the issue of how to provide people with the opportunity to find out about potentially beneficial, but novel technologies, such as companion robots, form opinions as to their likely value, and engender attitudes that are grounded in real world examples. We show that research on the psychology of intergroup relations can be used to investigate and understand how to approach this challenge. The findings of the studies suggest that direct and extended contact can change people's attitudes towards social robots and, more broadly, that methods and approaches borrowed from research on human-human relations can be used to understand how people are likely to interact with robots now and in the future. As such, the findings are likely to be of interest to psychologists, academics working in the fields of robotics and assistive technology, as well as to scientists and commercial organizations responsible for the development and dissemination of such technologies.

Chapter VII

Discussion

7.1. Overview

This final chapter explains the goals and contributions of this thesis. It summarises the findings obtained in both the systematic review and the empirical work. It also considers limitations to the research carried out in this thesis and explores what future research can be done to either assess these limitations or carry out further investigations to continue with unanswered questions that arise as a result of the research reported in this thesis.

7.2. The goals of this thesis

The aim of this thesis was to analyse people's attitudes, acceptance, anxiety and trust towards robots and the factors that can affect them in our current society. Both the systematic review and the empirical research explore and investigate these questions in a different way. The review reported in Chapter 3 synthesised the outcomes of 97 studies and examined a wide range of factors that could affect people's attitudes, acceptance, anxiety and trust towards robots. In the experimental studies reported in Chapters 4, 5 and 6, a social psychology method was applied in a Human-Robot Interaction (HRI) setting; specifically, techniques borrowed from research on intergroup relations were employed, where the robot was understood as an out-group member. In the empirical research, both direct and extended contract were implemented as independent variables. Attitudes and trust towards robots were measured using pre-existing and validated methods developed by experts in the field of HRI. This chapter will discuss in depth the implications and contributions of the main findings of the work mentioned previously.

7.3. Outcomes of the systematic review

The findings of the systematic review suggested that people typically have slightly positive opinions on social robots. Although this contradicts the widely accepted belief that people perceive robots as being threatening or dangerous to use, it does not necessarily mean that people will start acquiring social robots to aid them in their daily life. That being said, even if, attitudes do not inevitably correlate with behaviour (Eagly & Chaiken, 1993; Forscher et al, 2019; Lemon, 1973), they are likely to indirectly affect behavioural outcomes and shape intentions. The Theory of Planned Behaviour (TPB), brought up by Icek Ajzen (1991), states that "[a]ttitudes towards the behavior, subjective norms with respect to the behavior, and perceived control over the behavior are usually found to predict behavioral intentions with a high degree of accuracy" (Ajzen, 1991, pp. 206). So, it is possible to use attitudes as indirect predictors of behaviour but it is also necessary to take into account that they might not be the only factor affecting behavioural intentions.

It is a fact that people are not introducing robotic technologies in their daily life even if they do not have negative attitudes towards them. This suggests that there could be many other reasons for this to happen. Little research has been done investigating this topic. It may be that the present market does not cater the need of a non-expert private consumer. Maybe robots are too expensive or maybe they are not commercially easily accessible. It is also possible that people do not need such technology in their daily life or they do not see how their life would improve by doing it. This contrasts with, for example, the mobile phone, which clearly fulfilled the consumer's need to be connected with other people without having locational constraints.

Another technology that would be closer to the field of robotics would be the autonomous vacuum cleaner or the multi-cooker, which made tedious house chores much easier and comfortable improving the daily life of their consumers. Social robotics has the potential to solve real life problems and improve our current society (Dautenhahn, 2007a; Goodrich, & Schultz, 2008; Hans, et al., 2002; Harmo, et al., 2005; Kachouie, et al., 2014; Roy et al., 2000; Scopelliti, et al., 2005). That being said, there is a need for research investigating how social robots would contribute to that, which will be discussed later in this chapter in the section "Future work".

7.4. Outcomes of the empirical research

The general outcome of the empirical studies is that both direct and extended contact with robots affect positively attitudes towards robots. Taking these results into account and also considering similar studies in the field, it can be concluded that social psychology methods used to study intergroup relationships can actually be used in the field of HRI seeing robots as out-group members. This could be related to the media equation and the fact that people are able to see humanoid robots as social entities instead of lifeless machines.

The studies reported in this thesis found that participants had more positive implicit attitudes towards robots after extended contact. Specifically, participants (who did not meet a robot themselves) evidenced more positive implicit attitudes after watching a video of a friend describing their interaction with a robot. However, contrary to our expectations, we found no evidence that extended contact influenced participants' explicit attitudes. Previous studies suggest that discordance between implicit and explicit measures is common (Echabe, 2013). For example, some studies show that explicit attitudes can be changed more easily than implicit attitudes (Gawronski & Strack, 2004; Gregg et al., 2006; Petty et al., 2006), while others have found the opposite (Barden et al., 2004; Dasgupta & Greenwald, 2001; Wittenbrink et al., 2001). Nevertheless, these studies agree in one aspect; that there are situations in which implicit attitudes can change while explicit attitudes remain unchanged (for a review, see Forscher et al., 2019).

One situation when explicit and implicit attitudes can diverge is when stimuli are presented outside conscious awareness (Rydell et al., 2006); however, the extended contact procedures that we implemented were relatively explicit, suggesting that it is unlikely that participants were unaware of being exposed to attitude-relevant information. However, although participants may be aware of a stimulus, they may be unaware of the way in which it has been interpreted and / or influences their responses (Bargh & Chartrand, 2000). It is therefore possible that, although participants were aware that they had watched a video of a friend talking about the robot, they were not aware of the effect that this had on their attitudes – in other words, extended contact served as a 'supraliminal' priming procedure (Bargh & Chartrand, 2000). Indeed, previous research has shown that implicit attitudes can be shaped by recent experiences (Dasgupta & Greenwald, 2001; Lowery et al., 2001; Mitchell et al., 2003) and that these changes may occur outside of explicit awareness or conscious control.

An alternative explanation is that participants interpreted extended contact (but not direct contact) as intended to change their opinion and resisted this potential influence. For example, a meta-analysis of the effects of selective exposure to information (Hart et al., 2009) suggested that, in some cases, people do not change their opinion even if they have evidence that challenges their beliefs. It is therefore possible that participants did not want to change their attitudes only by hearing a friend's opinion, but were more inclined to change if they had experienced the interaction with the robot by themselves; perhaps because they did not view this direct interaction as intended to change their attitudes.

Moving forward, one might think that some mediators could be affecting the results obtained in the empirical studies. For example, it is reasonable to think that a higher score in the IOS or a higher similarity in the IOS between friendship pairs could predict stronger extended contact effect. In other words, one would think that a closer friendship would have a greater impact in the EC condition. Nonetheless, there was no correlation between these variables. Zhou, Page-Gould, Aron, Moyer and Hewstone (2019) carried out a meta-analysis investigating the factors affecting extended contact and their results suggest that “extended contact works to a similar degree regardless of how close one is to the in-group member or how close the in-group member is to the cross-group friend.” (Zhou et al., 2019, p. 153) This would explain the lack of correlation between the IOS and any dependent variable in the empirical studies; it appears that friendship does not moderate extended contact.

7.5. Contributions

First, if social robots are going to be more present in the future, then there is a need to understand how people think and feel about them. There was lack of evidence on attitudes towards social robots since previous systematic reviews or meta-analysis either did not focus specifically on this kind of robots or they centred their attention on a particular use of social robots (e.g., healthcare or education), instead of giving a bigger picture on attitudes towards social robots. In addition, previous research did not combine different measures (attitudes, anxiety, acceptance and trust) in order to obtain the valence (positive, negative) of the outcomes. Thus, the systematic review provided in this thesis, which synthesises this information that was missing previously, provides information about the people’s attitudes valence in a compact and efficient way. The findings indicate that, in general, people have a slightly positive attitude towards social robots, contradicting the belief that people perceive robots in a negative way. These findings will be useful because it can make researchers rethink the reasons why the use of social robotics is not widely spread and aim their research accordingly, keeping in mind that the reason this is happening is not because people see robots negatively.

The systematic review reported in Chapter 3 also identified which factors affect attitudes towards robots. The review did not find enough empirical evidence to conclude if the design of the robot and the application affected outcomes. Previous research claims age to be one of the factors affecting attitudes towards robots (Kuo et al., 2009), however, in the systematic review, age was not a significant moderator. In contrast, gender appeared to be one of the factors affecting trust in robots, with females typically being more trusting in robots. In addition, the type of exposure to the robot (direct contact, indirect contact or no contact) was the factor that had the strongest association with attitudes. This information provides a framework of knowledge that can be developed further by future research, which is discussed later in this chapter.

The systematic review also developed a new method for standardising the measures of participants' attitudes towards robots in each of the primary studies. This made it possible to combine all the measures and create a score that would tell us the valence of the outcomes, enabling homogeneity between the outcomes of studies and, hence, the possibility of analysing across them. A traditional meta-analysis was not viable since the review analysed people's attitudes, anxiety, trust and acceptance towards robots in general instead of measuring the effects of an intervention. That is, calculating the effect size of independent variables in each study was not the purpose of the review. Traditional effect size metrics could not be used to describe the valence of people's attitudes. So, an alternative method of standardising and analysing the data that did not rely on effect sizes was used. The potential approach of this method is to analyse a dependent variable (in this case, attitudes towards robots) without focusing on a specific independent variable affecting the former. Instead, this method is based on a more general approach; each factor that could affect the dependent variable is analysed as a moderator. Thus, this new method could open the doors to a new way of carrying out systematic reviews and meta-analysis which is useful when traditional methods cannot be applied. While this method has its limitations, it does allow for some statistical analysis of the literature and, hence, offers a better understanding of people's attitudes towards robots and the factors that influence those attitudes.

The empirical research carried out in this thesis helps to answer questions that have been neglected by previous research using the psychology of intergroup relations to understand how people relate to social robots. Specifically, although previous research has investigated the effects of different types of contact on people's attitudes towards robots (Conti, Di Nuovo, Buono, & Di Nuovo, 2017; Wullenkord, & Eyssel, 2014; Zlotowski et al., 2015), studies have usually focused on direct, mediated and imagined contact and have not yet considered the effects of extended contact. In contrast, the studies reported in this thesis investigated how both direct and extended contact affected both implicit and explicit measures of attitude. The finding that direct contact affected explicit attitudes is congruent with the studies done in the past (e.g., Wullenkord et al., 2016; Zlotowski et al., 2015) and supports previous knowledge. The results obtained regarding extended contact, however, add new information and suggest that extended contact might be used to promote positive relationships between people and social robots. However, the findings of the studies reported in this thesis suggest that extended contact only affected participants' implicit attitudes while it did not influence explicit attitudes. This was unexpected since the hypothesis stated that it would affect both implicit and explicit measures. However, these findings, while not predicted, do expand our understanding of how implicit and explicit measures can be affected by different intergroup contact procedures. Moreover, they highlight the importance of taking both types of attitudes into account when carrying out studies because they may measure different dimensions of attitudes and may be they influenced by different mechanisms.

In addition, the results obtained in the studies reported in this thesis indicate that extended contact can be used as a way to make people gain experience and knowledge about robots (for example, when other types of contact are not possible). Previous studies on human participants already showed that extended contact is effective in reducing intergroup tensions and hostility (Lemmer, & Wagner, 2015). A practical example would be the one presented in the study by Cameron, Rutland, Brown, and Douch (2006) in which children read stories about intergroup friendships, which leads them to have more positive attitudes toward refugees. The studies presented in this thesis suggest that extended contact with robots also has an effect on attitudes towards robots. This is relevant because extended contact could be used as a substitute to other types of contact that require more effort and resources. Extended contact requires minimal or no equipment, is affordable and can be done in many contexts and settings, which makes it a good alternative to other forms of contact.

The present thesis proposed extended contact as a way to provide a realistic reference of a social humanoid robot (opposed to fictional robots). As stated previously, direct contact has been used repeatedly in HRI and, consequently, in this aspect, this thesis only contributed by giving another example of a realistic reference such as a real humanoid robot. That is to say, participants who had direct contact with the robot had this accurate reference. Apart from that, by using extended contact, the research described in this thesis suggests a new way of providing a realistic reference of robots that could influence people's attitudes. Participants who saw their friend talking about the robot based their attitudes on this account. This is a realistic reference in the sense that it was portraying a real robot; participants in the video were not talking about a fictional robot but a real robot which had a real interaction with them. It is worth remembering that extended contact with robots is knowing someone who had interacted with a robot previously. This means that it is impossible to have extended contact with a fictional robot because it is impossible to interact with fictional robots (i.e. robots that do not exist) and, as a consequence, it is impossible to know someone who has interacted with a fictional robot. As a result, extended contact with robots will always be based upon a real robot and, therefore, provide a realistic reference.

If someone wanted to introduce robotics in a care centre or a company, they could start with implementing extended contact to provide a realistic reference of what a robot is and how it works. The empirical work provides extended contact as a way to make people gain knowledge about robots. Previous studies with human participants already show that extended contact is effective reducing intergroup tension. This is relevant because extended contact could be implemented when other types of contact are not possible. Extended contact is cheap, it requires minimal equipment, it can be done in many different places and it provokes less anxiety because of personal detachment.

The empirical research expanded on existing knowledge in the field of HRI since it provided more evidence to suggest that techniques developed in the social sciences can be used to study how people interact with robots. Many previous studies have applied these methods successfully, which could be related to the fact that people are able to see some objects (including robots) as social agents and not just as a tool. This thesis applied inter-group research methodologies that are normally used to study prejudice towards minorities or different ethnic groups. Taking all the precedents into account and how this method worked in this thesis, one can conclude that the use of such techniques assists in the development of HRI and facilitates a better comprehension about how people perceive social robots.

7.6. Limitations

There are a number of limitations to the research reported in the thesis, which are worth identifying and considering, especially with a view to how they might be addressed in future work. First, all the studies in the systematic review were classified according to the factor that could potentially affect attitudes towards robots (e.g., the design of the robot, type of exposure to the robot, domain of application and sample characteristics). One of the main issues with this method is that, sometimes, there were not enough studies in some categories. For example, only 4 studies analysed attitudes towards a social zoomorphic robot. Therefore, the category of zoomorphic robots had to be merged with other categories so that these primary studies could be included in the review. A similar situation happened with androids or robots for domestic use. It would be useful for future meta-analysis and systematic reviews to have enough primary studies in order to classify them into more specific categories and, in this way, be able to provide a more specialized and fragmented analysis. Thus, it would be helpful to encourage researchers to perform studies that would expand those minority categories.

The main limitation in the empirical research is the fact that attitudes were measured right before and after the interventions. Consequently, there was no way to know if the effects lasted in time. For this reason, longitudinal studies can contribute greatly in exploring methods to influence attitudes and if these effects last in time. For example, Gawronski et al. (2017) carried out a series studies in which they investigated the temporal stability of both implicit and explicit attitudes. Their results suggest that implicit measures have less temporal stability than explicit measures. However, there is no further research validating this results. So, as promising as they might seem, it is essential to proceed with caution. Another option, without carrying out a longitudinal study, could be measuring the participants' attitudes waiting a certain period of time after the intervention, for example, one hour later.

Another thing that should be considered regarding the empirical studies is the fact that extended contact with the robot was given in a video format. That is to say, participants watched a video of their friend talking about the robot. It would be interesting to see if extended contact where the two participants talk face to face would have had a different effect on attitudes. In addition to what has being said, this thesis gives some explanations that try to rationalize the discrepancy between implicit and explicit attitudes in participants who were exposed to extended contact with robots. Nonetheless, these explanations are mere speculations. So, as a final suggestion, it would be useful to investigate further about this discrepancy and the causes of this phenomenon. For example, it would be helpful to know if participants were affected by a confirmation bias or if it was a priming effect.

7.7. Future work

This section is aimed at discussing the future work that could be done in order to explore or expand the questions that this thesis might cause or leave unsolved. Starting with the questions that may arise from the systematic review, future research could analyse why most people are not introducing social robotics in their daily life. The results obtained in the systematic review indicate that people do not have negative attitudes towards robots, therefore, this phenomenon could not be assessed as a matter of attitudes. In section 7.3. this thesis speculates by giving some possible reasons this might be happening. It would be useful if researchers did a survey specifically asking about this particular issue. It could be done by asking a question with an open answer, for instance "Why are you not introducing social robots in your life?", and then perform a qualitative analysis by examining the most common reasons participants would have written. Another approach could consist of administering a questionnaire, asking the same question, in which participants would have several check boxes representing different reasons they are not using social robots (e.g. "I do not trust robots", "robots are difficult to use", "robots are expensive", "robots are not available in shops", "robots would not improve my life", "robots are creepy" etc.). It would be helpful if participants could check more than one box since it is possible that they have more than one reason that is keeping them from introducing this technology in their lives. After that, the analysis could consist on counting the frequency in which the boxes are checked. Up until now, there is no research analysing this subject, which would provide an idea of the most common reasons why social robotics is not being introduced in people's daily lives.

To continue, some previous research suggest that age is a significant variable that affects the acceptance of robots (Kuo et al., 2009) while other research has inconclusive results (De Graaf, & Allouch, 2013a). The systematic review in this thesis suggests that age is not an influential factor. However, in order to investigate this, the review used the mean age of the participants of each study. This lack of granularity could have affected the outcomes of the review. In this sense, because there is no conclusive empirical evidence proving that age is a decisive factor in attitudes towards social robots, future research needs to address this topic and investigate age as a sole factor. For instance, participants from all ages could complete several tasks measuring their attitudes towards robots, and after that, a regression analysis could determine if there is an actual correlation between age and the scores obtained in the measurements. Maybe this would finally shed some light on the topic of age being a factor affecting attitudes towards robots.

Now addressing the questions that may be arisen by the empirical work, there is the obvious question of extended contact affecting implicit attitudes but not explicit attitudes. It is possible for future research to investigate if it is a case of supraliminal priming only. A study like this would follow a very similar protocol to the one used in Study 4 with some modifications. First of all, there would be no interaction with any robot. Participants would come in pairs of friends and they would be separated in two conditions. In the first condition (which would be similar to the control condition presented in Study 4), participants would complete all the measures and, once all data would have been collected, they would record a video narrating a text or saying words related to social robotics. In the second condition (which would be similar to the experimental condition in Study 4), participants would watch this video. Measurements would be taken before and after the video and then compared doing the same analysis that was done in this thesis. It is important to notice that the person in the video would not say that they have interacted with a robot and, for this reason, in this study, extended contact would not be present. If the results replicated the ones obtained in this thesis, it would be likely that this is a case of supraliminal priming only and, therefore, extended contact would not have been the factor affecting the results.

Apart from this, it is also possible to test if the findings of the empirical work are a result of a defensive reaction. This kind of study would have one control condition and two experimental conditions, namely Positive Bias (PB) condition and Negative Bias (NB) condition. Participants would come again in pairs of friends. Each pair would have one participant in the control condition, the other participant would go randomly either to the PB or NB condition, which will determine what kind of video they would see. The participant in the control condition would complete all the measurements and, after that, they would record a video talking positively (if their friend is in the PB condition) or negatively (if their friend is in the NB condition) about social robots. In the experimental conditions, participants would watch the video their friend would have had recorded. As always, measurements would be completed before and after watching the video. The analysis would test if participants in the PB condition changed their attitudes in a negative way (showing a defensive reaction towards the video) and if participants in the NB condition changed their attitudes in a positive way. If results showed a change in implicit but not explicit measures, it would mean that, in the empirical work carried out in this thesis, participants just had a defensive reaction and that, maybe, extended contact was not the only factor affecting their attitudes.

Other interesting research could analyse attitudes towards robots in a longitudinal study in which participants would live together with a social robot. It would be interesting to see if long-term interactions would affect their attitudes towards robots and in what way. It would be appealing to perform this experiment with a social humanoid robot. However, depending on the circumstances, resources might be limited and, therefore, it would be unrealistic to do so. Alternatively, this study could be carried out with pet robots. Explicit and implicit measurements could be taken periodically in order to see if there is an evolution.

In a similar way, coexisting with robots in the same house would be useful in order to measure behaviour towards robots and how it correlates with behavioural intentions. As mentioned previously in this thesis, the TPB (Ajzen, 1991) suggests that intentions and attitudes can have an influence and shape behaviour. Therefore, it would be valuable to know how attitudes towards social robots affects people's actual behaviour. Although social robots are more commercially available now than a decade ago, they are still not a big part of our daily lives. Therefore, it would be challenging to perform an experiment measuring actual behaviour towards robots. One of the possible implementations for an experiment of this nature could be done by giving one social robot to each of the participants in the study. This robot would then coexist with the participant in the same house for a certain period of time. One of the measures that could be taken in order to measure the participant's behaviour towards the social robot is the time that they spend interacting with the robot. Participants would also perform tasks or questionnaires measuring their attitudes and then researchers could analyse how these two variables interact with each other.

In a more general view, future work could involve research adapting other techniques used in intergroup or social psychology studies and applying them in the field of HRI. For example, one of the most famous experiments about intergroup conflict is "The Robbers Cave" experiment, carried out by Muzafer Sherif in the 50s (Sherif, 2010). This study took place at a holyday camp in the Robbers Cave Park in Oklahoma and it involved three phases. The first phase took place during the first week of their stay; children between the ages of 11 and 12 were divided (without them knowing it) into two different groups. Then these two different groups took part in activities so they could bond among their in-group members. After a week, in the second phase, the children were told about the existence of the other group. At that point, these two groups were asked to perform competitive games, which aroused intergroup conflict. In the third phase, the same children were then asked to perform cooperative games which involved working together to achieve a certain goal. This help the children to bond with the out-group members and eventually led to the dissolution of any intergroup conflict.

It would be interesting to know how this experiment would differ if one of the groups were composed by humanoid robots. This would require a huge amount of resources since researchers would need to have access to a big number of humanoid robots that would be able to interact with humans smoothly and naturally. Implementing a Wizard of Oz setting would be also a possibility but researchers would still need to have several robots and many "wizards" to control them. One of the biggest differences would be that in the original experiment, the outgroups were created randomly and artificially because, after all, they were all children from a similar background. In a HRI setting, the robots would be clearly perceived as an out-group without the intervention of the researchers. This is because robots already look and behave different than humans and, therefore, it is easier to perceive them as out-group members. This could affect the results and make them different from the original “Robbers Cave” experiment. Despite the huge amount of resources needed for this research, it would be helpful to know how intergroup dynamics work in a larger scale with a bigger number of individuals.

It would be impossible to give a specific example of each social psychology experiment and how to adapt it to a HRI setting. Nonetheless, this thesis would like to encourage future researchers to take a closer look to these methods and see the value in them. Social psychology techniques work in the HRI field because humans interact with robots in a similar way as they interact with other people. For this reason, adapting these methods from the psychological field can prove very useful in learning interrelation dynamics between robots and humans.

7.8. Conclusions

To sum up, the aim of this thesis was to explore how people perceive social robots and the factors that affect their views. Through a systematic review that analysed people’s attitudes, anxiety, acceptance and trust in robots, this thesis provided new information and clarifications on which factors affect how people see social robots. This was done using an innovative method to calculate a standardized score that allowed the analysis across the studies, which used different methods and scales. The findings of this review showed that people have slightly positive attitudes towards robots and the main factor affecting such was the type of exposure to the robots (direct or indirect exposure) that participants had in different studies. Empirical research investigating the effect of direct and extended contact in people’s explicit and implicit attitudes, then contributed with new data and findings that addressed key unanswered questions in previous investigations. The results supported previous studies because they indicated that direct contact affected both explicit and implicit attitudes. In contrast, extended contact affected only implicit attitudes towards robots while explicit attitudes remained the same, which was not anticipated. This thesis explores different facts that could justify these findings and provides different rationalizations. Finally, the field of HRI is still a new area of investigation and there are many neglected points that need further investigation. This thesis provides new ideas for future research that would contribute to the study of how people interact with robots and would shed some light to some of the issues discussed previously.

References:

* Papers marked with an asterisk were included in the systematic review.

ACM/IEEE International Conference on Human-Robot Interaction. (2020). *About*. Retrieved January 25, 2020 from https://humanrobotinteraction.org/2020/about/

Active-Robots.com. (2014). *Robots found in the classroom*. Retrieved October 1, 2019 from https://www.active-robots.com/our-blog/robots-found-in-the-classroom/

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211. doi: 10.1016/0749-5978(91)90020-T

Aldebaran. (n.d.). *Pepper, the humanoid robot from Aldebaran, a genuine companion*. Retrieved February 27, 2017 from https://www.ald.softbankrobotics.com/en/cool-robots/pepper

*[44] Alemi, M., Meghdari, A., & Ghazisaedy, M. (2014). The effect of employing humanoid robots for teaching English on students' anxiety and attitude. In *2014 Second RSI/ISM International Conference on Robotics and Mechatronics, IEEE*, 754-759. doi: 10.1109/ICRoM.2014.6990994

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A Handbook of Social Psychology*, (pp. 789-844). Clark University Press.

Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

American National Standards Institute (1999). *RIA/ANSI R15.06 1999 American national standard for industrial robots and robot systems safety requirements*. American National Standards Institute, New York

Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological bulletin, 71*(5), 319. doi: 10.1037/h0027352

Anki Vector: The Home Robot With interactive AI Technology. (2019). Retrieved January 27, 2020 from https://anki.com/en-gb/vector.html

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology, 63*(4), 596. doi: 10.1037/0022-3514.63.4.596

Asif, M., Sabeel, M., & Mujeeb-ur-Rahman, K. Z. (2015). Waiter robot–solution to restaurant automation. In *Proceedings of the 1st student multi disciplinary research conference (MDSRC), At Wah, Pakistan*, 14-15.

*[91] Avelino, J., Moreno, P., Bernardino, A., Correia, F., Paiva, A., Catarino, J., & Ribeiro, P. (2018). The Power of a Hand-shake in Human-Robot Interactions. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain*, 1864-1869. doi:10.1109/IROS.2018.8593980

*[70] Backonja, U., Hall, A. K., Painter, I., Kneale, L., Lazar, A., Cakmak, M., Thompson, H. J., & Demiris, G. (2018). Comfort and Attitudes Towards Robots Among Young, Middle-Aged, and Older Adults: A Cross-Sectional Study. *Journal of Nursing Scholarship, 50*(6), 623-633. doi: 10.1111/jnu.12430

Bandura, A. (2009). Social cognitive theory of mass communication. In J. Bryant, & M. B. Oliver (Eds.), *Media effects* (pp. 110-140). Routledge.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: the impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, *87*(1), 5-22. doi: 10.1037/0022-3514.87.1.5

Bargh, J. A., & Chartrand, T. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253-285). Cambridge University Press.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: the influence of trait information presented outside of conscious awareness on impression formation. *Journal of personality and Social psychology, 43*(3), 437-449. doi: 10.1037/0022-3514.43.3.437

*[92] Barnes, J., Richie, E., Lin, Q., Jeon, M., & Park, C. H. (2018). Emotive Voice Acceptance in Human-Robot Interaction. In *The 24th International Conference on Auditory Display (ICAD).*

Bartneck, C., Chioke, R., Menges, R., & Deckers, I. (2005a). Robot Abuse – A Limitation of the Media Equation. In *Proceedings of the Interact 2005 Workshop on Abuse, Rome*.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics, 1*(1), 71-81.

*[52] Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2005). Cultural differences in attitudes towards robots. In *AISB Symposium on Robot Companions: Hard Problems And Open Challenges In Human-Robot Interaction, Hatfield*, 1-4

Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., & Kennsuke, K. (2005b). A cross-cultural study on attitudes towards robots. In *HCI international, Las Vegas*.

*[51] Bartneck, C., Suzuki, T., Kanda, T., & Nomura, T. (2007a). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *Ai & Society, 21*(1-2), 217-230.

Bartneck, C., Verbunt, M., Mubin, O., & Al Mahmud, A. (2007b). To kill a mockingbird robot. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction,* 81-87. doi: 10.1145/1228716.1228728

*[80] Bazzano, F., & Lamberti, F. (2018). Human-Robot Interfaces for Interactive Receptionist Systems and Wayfinding Applications. *Robotics, 7*(3), 56. doi: 10.3390/robotics7030056

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics, 3*(21), eaat5954. doi: 10.1126/scirobotics.aat5954

*[53] Bernotat, J., Eyssel, F. (2018). Can('t) Wait to Have a Robot at Home? - Japanese and German Users' Attitudes Toward Service Robots in Smart Homes. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),* 15-22. doi:10.1109/ROMAN.2018.8525659

Bhattacherjee, A., Sanford, C. (2009). The intention–behaviour gap in technology usage: the moderating role of attitude strength. *Behaviour & Information Technology, 28*(4), 389-401. doi: 10.1080/01449290802121230

Birk, A. (1998). Behavior-based robotics, its scope and its prospects. In *IECON'98. Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society (Cat. No. 98CH36200), Vol. 4* (pp. 2180-2185). IEEE. doi: 10.1109/IECON.1998.724059

*[93] Birnbaum, G. E., Mizrahi, M., Hoffman, G., Reis, H. T., Finkel, E. J., & Sass, O. (2016). Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 165-172. doi: 101109/HRI20167451748

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81,*828 – 841. doi: 10.1037/0022-3514.81.5.828

Bongaarts, J. (2009). Human population growth and the demographic transition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1532), 2985-2990. doi: 10.1098/rstb.2009.0137

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies, 59*(1), 119-155. doi: 10.1016/S1071-5819(03)00018-1

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of personality and social psychology, 47*(6), 1191-1205. doi: 10.1037/0022-3514.47.6.1191

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual review of psychology, 68,* 627-652. doi: 10.1146/annurev-psych-010416-043958

Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of healthcare robots for the older population: Review and future directions. *International journal of social robotics, 1*(4), 319-330. doi: 10.1007/s12369-009-0030-6

*[94] Broadbent, E., Kerse, N., Peri, K., Robinson, H., Jayawardena, C., Kuo, T., Datta, C., Stafford, R., Butler, H., Jawalkar, P., Amor, M., Robins, B., & MacDonald, B. (2016). Benefits and problems of health-care robots in aged care settings: A comparison trial. *Australasian journal on ageing, 35*(1), 23-29. doi: 10.1111/ajag.12190

*[2] Broadbent, E., Tamagawa, R., Patience, A., Knock, B., Kerse, N., Day, K., & MacDonald, B. A. (2012). Attitudes towards health-care robots in a retirement village. *Australasian journal on ageing, 31*(2), 115-120. doi: 10.1111/j.1741-6612.2011.00551.x

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: a review. *Gerontechnology*, *8*(2), 94-103. doi:10.4017/gt.2009.08.02.002.00

*[134] Brule, R., Dotsch, R., Bijlstra, G., Wigboldus, D. H., & Haselager, P. (2014). Do Robot Performance and Behavioral Style affect Human Trust?. *International Journal of Social Robotics, 4*(6), 519-531. doi:10.1007/s12369-014-0231-5

Burns, J. (2012, November 8). *Robots in the classroom help autistic children learn*. BBC. https://www.bbc.co.uk/news/education-20252593

*[95] Cameron, D., Millings, A., Fernando, S., Collins, E. C., Moore, R., Sharkey, A., & Prescott, T. (2018). The effects of robot facial emotional expressions and gender on child–robot interaction in a field study. *Connection Science, 30*(4), 343-361. doi:101080/0954009120181454889

Cameron, L., Rutland, A., Brown, R., & Douch, R. (2006). Changing children's intergroup attitudes toward refugees: Testing different models of extended contact. *Child development, 77*(5), 1208-1219. doi: 10.1111/j.1467-8624.2006.00929.x

*[45] Carrillo, F. M., Butchart, J., Kruse, N., Scheinberg, A., Wise, L., & McCarthy, C. (2018). Physiotherapists' Acceptance of a Socially Assistive Robot in Ongoing Clinical Deployment. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2018. IEEE,* 850-855. doi: 10.1109/ROMAN.2018.8525508

Carroll, J. (2014). Human Computer Interaction - brief intro. In K. Dautenhahn, & C. Ghauoi (Eds.), *The Encyclopedia of Human-Computer Interaction*, (2nd Ed.) Interaction Design Foundation.

*[96] Carter, E. J., Mistry, M. N., Carr, G. P. K., Kelly, B. A., & Hodgins, J. K. (2014). Playing catch with robots: Incorporating social gestures into physical interactions. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinbrugh, UK*, 231-236. doi:10.1109/ROMAN.2014.6926258

*[97] Cavallo, F., Esposito, R., Limosani, R., Manzi, A., Bevilacqua, R., Felici, E., & Dario, P. (2018). Acceptance of Robot-Era system: results of robotic services in smart environments with older adults. *Journal of Medical Internet Research*.

*[98] Chang, R. C. S., Lu, H. P., & Yang, P. (2018). Stereotypes or golden rules? Exploring likable voice traits of social robots as active aging companions for tech-savvy baby boomers in Taiwan. *Computers in Human Behavior, 84,* 194-210. doi:101016/jchb201802025

*[46] Chang, W-L., White, J. P., Park, J., Holm, A., & Šabanović, S. (2012). The effect of group size on people's attitudes and cooperative behaviors toward robots in interactive gameplay. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, 2012. IEEE*, 845-850. doi: 10.1109/ROMAN.2012.6343857

Chen, K., & Chan, A. H. (2011). A review of technology acceptance by older adults. *Gerontechnology, 10*(1), 1-12. doi: 10.4017/gt.2011.10.01.006.00

Chen, T. L., Bhattacharjee, T., Beer, J. M., Ting, L. H., Hackney, M. E., Rogers, W. A., & Kemp, C. C. (2017). Older adults' acceptance of a robot for partner dance-based exercise. *PloS one, 12*(10). doi: 10.1371/journal.pone.0182736

Cheney, M., & Uth, R. (1999). *Tesla, master of lightning*. Barnes & Noble Publishing.

Christakis, N. (2019, April). *How AI will rewire us*. The Atlantic. https://www.theatlantic.com/

*[99] Coco, K. P. R. N., Kangasniemi, M. P., & Rantanen, T. P. (2018). Care Personnel's Attitudes and Fears Toward Care Robots in Elderly Care: A Comparison of Data from the Care Personnel in Finland and Japan. *Journal of Nursing Scholarship, 50*(6), 634-644. doi: 10.1111/jnu.12435

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), New Jersey: Lawrence Erlbaum Associates

*[100] Conti, D., Cattani, A., Di Nuovo, S., & Di Nuovo, A. (2015). A cross-cultural study of acceptance and use of robotics by future psychology practitioners. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 555-560. doi:101109/ROMAN20157333601

*[58] Conti, D., Di Nuovo, S., Buono, S., & Di Nuovo, A. (2017). Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals. *International Journal of Social Robotics, 9*(1), 51-62. doi:101007/s12369-016-0359-6

Cook, S. W. (1962). The systematic analysis of socially significant events: A strategy for social research. *Journal of Social Issues, 18*(2), 66-84. doi: 10.1111/j.1540-4560.1962.tb02201.x

Cook, S. W., & Selltiz, C. (1955). Some factors which influence the attitudinal outcomes of personal contact. *International Social Science Bulletin, 7*, 51-58.

Copeland, B. J. (2000). The turing test. *Minds and Machines, 10*(4), 519-539. doi: doi.org/10.1023/A:1011285919106

*[101] Correia, F., Guerra, C., Mascarenhas, S., Melo, F. S., & Paiva, A. (2018). Exploring the impact of fault justification in human-robot trust. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 507-513.

Costin, F. (1969). The scrambled sentence test: A group measure of hostility. *Educational and Psychological Measurement, 29*(2), 461-468. doi: 10.1177/001316446902900224

*[67] Cramer, H., Kemper, N., Amin, A., Wielinga, B., & Evers, V. (2009). 'Give me a hug': the effects of touch and autonomy on people's responses to embodied social agents. *Computer Animation and Virtual Worlds, 20*(2-3), 437-445. doi: 10.1002/cav.317

Crisp, R. J., & Turner, R. N. (2009). Can imagined interactions produce positive perceptions?: Reducing prejudice through simulated social contact. *American Psychologist*, *64*(4), 231-240. doi: 10.1037/a0014718

Crisp, R. J., & Turner, R. N. (2012). Imagined intergroup contact: Refinements, debates and clarifications. In G. Hodson & M. Hewstone (Eds.), *Advances in Intergroup Contact (pp. 135-151)*. Hove: Psychology Press (Taylor & Francis).

Crites Jr, S. L., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: Conceptual and methodological issues. *Personality and Social Psychology Bulletin*, *20*(6), 619-634. doi: 10.1177/0146167294206001

Crunchbase (2009). *Aldebaran Robotics acquired by SoftBank Robotics Holdings*. https://www.crunchbase.com/acquisition/softbank-robotics-acquires-aldebaran-robotics--9a05bfc7#section-overview

*[102] Damholdt, M. F., Nørskov, M., Yamazaki, R., Hakli, R., Hansen, C. V., Vestergaard, C., & Seibt, J. (2015). Attitudinal change in elderly citizens toward social robots: the role of personality traits and beliefs about robot functionality. *Frontiers in psychology, 6*, 1701. doi:103389/fpsyg201501701

Dario, P., Verschure, P. F., Prescott, T., Cheng, G., Sandini, G., Cingolani, R., Dillmanng, R., Floreanoh, D., Lerouxi, C., MacNeill, S., Roelfsemam, P., Verykiosn, X., Bicchio, A., Melhuishp, C., & Albu-Schäfferq, A. (2011). Robot companions for citizens. *Procedia Computer Science*, *7*, 47-51. doi: 10.1016/j.procs.2011.12.017

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*(5), 800-814. doi: 10.1037/0022-3514.81.5.800

Dautenhahn, K. (2007a). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical transactions of the royal society B: Biological sciences, 362*(1480), 679-704. doi: 10.1098/rstb.2006.2004

Dautenhahn, K. (2007b). Methodology & themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems, 4*(1), 15. doi: 10.5772/5702

Dautenhahn, K. (2014). Human-Robot Interaction. In K. Dautenhahn, & C. Ghauoi (Eds.), *The Encyclopedia of Human-Computer Interaction*, (2nd Ed.) Interaction Design Foundation.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly, 13*(3), 319-340. doi: 10.2307/249008

Dawe, J., Sutherland, C., Barco, A., & Broadbent, E. (2019). Can social robots help children in healthcare contexts? A scoping review. *BMJ paediatrics open, 3*(1). doi: 10.1136/bmjpo-2018-000371

DeCoster, J., & Claypool, H. M. (2004). A Meta-Analysis of Priming Effects on Impression Formation Supporting a General Model of Informational Biases. *Pers Soc Psychol Rev, 8*(1), 2-27. doi: 10.1207/S15327957PSPR0801_1.

Dediu, H. (2015). *Personal computer*. *Asymco*. http://www.asymco.com/2015/04/14/personal-computer

Desai, M., Stubbs, K., Steinfeld, A. & Yanco, H. (2009). Creating Trustworthy Robots: Lessons and inspirations from automated systems. In *Proceedings of the AISB Convention: New Frontiers in Human-Robot Interaction, Edinburgh, Scotland*. Retrieved from https://www.ri.cmu.edu/pub_files/2009/4/Desai_paper.pdf

*[57] Dinet, J., & Vivian, R. (2014). Exploratory investigation of attitudes towards assistive robots for future users. *Le travail humain, 77*(2), 105-125.

Dovidio, J. F., Eller, A., & Hewstone, M. (2011). Improving intergroup relations through direct, extended and other forms of indirect contact. *Group Processes & Intergroup Relations*, *14*(2), 147-160. doi: 10.1177/1368430210390555

*[50] Dunst, C. J., Trivette, C. M., Prior, J., Hamby, D. W., & Embler, D. (2013). Parents' Judgments of the Acceptability and Importance of Socially Interactive Robots for Intervening with Young Children with Disabilities. Social Robots Research Reports, Number 1. *Orelena Hawks Puckett Institute*

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Ebel, K. H. (1986). The impact of industrial robots on the world of work. *Int'l Lab. Rev., 125*(1), 39.

Echabe, A. E. (2013). Relationship between implicit and explicit measures of attitudes: The impact of application conditions. *Europe's Journal of Psychology*, *9*(2), 231-245. doi: 10.5964/ejop.v9i2.544

Eiser, J. (1986). *Social Psychology: Attitudes, Cognition and Social Behaviour.* Cambridge: Cambridge University Press.

Eiser, J. R., Miles, S., & Frewer, L. J. (2002). Trust, perceived risk, and attitudes toward food technologies1. *Journal of Applied Social Psychology, 32*(11), 2423-2433. doi: 10.1111/j.1559-1816.2002.tb01871.x

Eksiri, A., & Kimura, T. (2015). Restaurant service robots development in thailand and their real environment evaluation. *Journal of Robotics and Mechatronics, 27*(1), 91-102. doi: 10.20965/jrm.2015.p0091

Eller, A., & Abrams, D. (2004). Come together: Longitudinal comparisons of Pettigrew's reformulated intergroup contact model and the common ingroup identity model in Anglo-French and Mexican-American contexts. *European Journal of Social Psychology*, *34*(3), 229-256. doi: 10.1002/ejsp.194

Enz, S., Diruf, M., Spielhagen, C., Zoll, C., & Vargas, P. A. (2011). The social role of robots in the future—explorative measurement of hopes and fears. *International Journal of Social Robotics, 3*(3), 263. doi: 10.1007/s12369-011-0094-y

*[103] Erebak, S., & Turgut, T. (2019). Caregivers' attitudes toward potential robot coworkers in elder care. *Cognition, Technology & Work, 21*(2), 327-336. doi: 101007/s10111-018-0512-0

European Commission (2012, September) *Public Attitudes towards Robots*. https://www.ab.gov.tr/files/ardb/evt/Public_attitudes_toward_robots_2012.pdf

*[104] Eyssel, F., De Ruiter, L., Kuchenbrandt, D., Bobinger, S., & Hegel, F. (2012). 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 125-126. doi: 10.1145/2157689.2157717

*[105] Eyssel, F., Kuchenbrandt, D., Hegel, F., & de Ruiter, L. (2012). Activating elicited agent knowledge: How robot and user features shape the perception of social robots. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 851-857. doi: 10.1109/ROMAN.2012.6343858

*[106] Ezer, N., Fisk, A. D., & Rogers, W. A. (2009). Attitudinal and intentional acceptance of domestic robots by younger and older adults. In *International conference on universal access in human-computer interaction*, 39-48. doi:10.1007/978-3-642-02710-9_5

Fazio, R. H. (2001). On the automatic activation of associated evaluations: *An overview. Cognition & Emotion, 15*(2), 115-141. doi: 10.1080/02699930125908

Feng, H., Gutierrez, A., Zhang, J., & Mahoor, M. H. (2013). Can NAO robot improve eye-gaze attention of children with high functioning autism?. In *2013 IEEE International Conference on Healthcare Informatics*, 484-484. doi: 10.1109/ICHI.2013.72

*[107] Fischer, K., Weigelin, H. M., & Bodenhagen, L. (2018). Increasing trust in human–robot medical interactions: effects of transparency and adaptability. *Paladyn, Journal of Behavioral Robotics*, *9*(1), 95-109. doi:10.1515/pjbr-2018-0007

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intetion, and behavoir: An introduction to theory and research.* Reading, MA: Addison-Wesley

Fishbein, M. (1979). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation*, *27*, 65–116.

Forgas, J. P., Cooper, J., & Crano, W. D. (2011). *The psychology of attitudes and attitude change*. Psychology Press.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522-559. doi: 10.1037/pspa0000160

Foster, M. E., Alami, R., Gestranius, O., Lemon, O., Niemelä, M., Odobez, J. M., & Pandey, A. K. (2016). The MuMMER project: Engaging human-robot interaction in real-world public spaces. In *International Conference on Social Robotics*, 753-763. doi: 10.1007/978-3-319-47437-3_74

Freedy, A., DeVisser, E., Weltman, G., & Coeyman, N. (2007). Measurement of trust in human-robot collaboration. In *2007 International Symposium on Collaborative Technologies and Systems*, 106-114. doi: 10.1109/CTS.2007.4621745

Frey, D. (1986). Recent research on selective exposure to information. *Advances in Experimental Social Psychology*, *19*, 41-80. doi: 10.1016/S0065-2601(08)60212-9

*[62] Fridin, M., & Belokopytov, M. (2014). Acceptance of socially assistive humanoid robot by preschool and elementary school teachers. *Computers in Human Behavior*, *33*, 23-31. doi: 10.1016/j.chb.2013.12.016

*[17] Gaudiello, I., Zibetti, E., Lefort, S., Chetouani, M., & Ivaldi, S. (2016). Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers. *Computers in Human Behavior, 61*, 633-655. doi: 10.1016/j.chb.2016.03.057

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological bulletin, 132*(5), 692. doi: 10.1037/0033-2909.132.5.692

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40*(4), 535-542. doi: 10.1016/j.jesp.2003.10.005

Gawronski, B., Brochu, P. M., Sritharan, R., & Strack, F. (2012). Cognitive consistency in prejudice-related belief systems: Integrating old-fashioned, modern, aversive and implicit forms of prejudice. *Cognitive consistency: A fundamental principle in social cognition*, 369-389.

Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures: A longitudinal analysis. *Personality and Social Psychology Bulletin, 43*(3), 300-312. doi: 10.1177/0146167216684131

Gelsomini, M., Leonardi, G., Degiorgi, M., Garzotto, F., Penati, S., Silvestri, J., Ramuzat, N. & Clasadonte, F. (2017). Puffy-an inflatable mobile interactive companion for children with Neurodevelopmental Disorders. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2599-2606. doi: 10.1145/3027063.3053245

George, D. (2011). *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e.* Pearson Education India.

Geppert, L. (2004). Yoshihiro Kuroki: dancing with robots. *IEEE Spectrum, 41*(2), 34-35. doi: 10.1109/MSPEC.2004.1265131

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. *Computers in Human Behavior, 93,* 53-61. doi: 10.1016/j.chb.2018.11.0

*[18] Gombolay, M., Yang, X. J., Hayes, B., Seo, N., Liu, Z., Wadhwania, S., & Shah, J. (2018). Robotic assistance in the coordination of patient care. *The International Journal of Robotics Research, 37*(10), 1300-1316. doi: 101177/0278364918778344

Goodrich, M. A., & Schultz, A. C. (2008). Human–robot interaction: a survey. *Foundations and Trends in Human–Computer Interaction, 1*(3), 203-275. doi: 10.1561/1100000005

*[48] de Graaf, M. M., & Allouch, S. B. (2013a). Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems, 61*(12), 1476-1486. doi: 10.1016/j.robot.2013.07.007

de Graaf, M. M., & Allouch, S. B. (2013b). The relation between people's attitude and anxiety towards robots in human-robot interaction. In *2013 IEEE RO-MAN*, 632-637. doi: 10.1109/ROMAN.2013.6628419

*[37] de Graaf, M. M., Allouch, S. B., & Lutfi, S. (2016). What are people's associations of domestic robots?: Comparing implicit and explicit measures. *In 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN*, 1077-1083. doi: 10.1109/ROMAN.2016.7745242

de Graaf, M. M., Allouch, S. B., & van Dijk, J. A. (2019). Why would I use this in my home? A model of domestic social robot acceptance. *Human–Computer Interaction, 34*(2), 115-173. doi: 10.1080/07370024.2017.1312406

Greczek, J., Kaszubski, E., Atrash, A., & Matarić, M. (2014). Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders. In *The 23rd*

IEEE international symposium on robot and human interactive communication, 561-566. doi: 10.1109/ROMAN.2014.6926312

Green, A., Huttenrauch, H., & Eklundh, K. S. (2004). Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*, 575-580. doi: 10.1109/ROMAN.2004.1374824

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*(1), 4.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464-1480. doi: 10.1037//0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197-216. doi: 10.1037/0022-3514.85.2.197

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*(1), 1-20. doi: 10.1037/0022-3514.90.1.1

Groom, V., & Nass, C. (2007). Can robots be teammates? Benchmarks in human-robot teams. *Interaction Studies*, *8*(3), 483-500. doi:10.1075/is.8.3.10gro

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*(3), 1369. doi: 10.1037/a0035028

Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in physicians and medical institutions: what is it, can it be measured, and does it matter?. *The milbank quarterly, 79*(4), 613-639. doi: 10.1111/1468-0009.00223

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011a). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(5), 517-527. doi: 10.1177/0018720811417254

Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011b). Can you trust your robot?. *Ergonomics in Design, 19*(3), 24-29. doi: 10.1177/1064804611415045

Hans, M., Graf, B., & Schraft, R. (2002). Robotic home assistant care-o-bot: Past-present-future. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, 380-385. doi: 10.1109/ROMAN.2002.1045652

Harding, J., Kutner, B.M., Proshansky, H., & Chein, I. (1954). Prejudice and ethnic relations. In G. Lindzey (Ed.), *Handbook of social psychology.* Cambridge, Mass: Addison-Wesley.

*[108] Haring, K. S., Silvera-Tawil, D., Matsumoto, Y., Velonaki, M., & Watanabe, K. (2014). Perception of an android robot in Japan and Australia: A cross-cultural comparison. In *International conference on social robotics* (pp. 166-175). Springer, Cham. doi:10.1007/978-3-319-11973-1_17

*[109] Haring, K. S., Silvera-Tawil, D., Takahashi, T., Velonaki, M., & Watanabe, K. (2015). Perception of a humanoid robot: a cross-cultural comparison. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 821-826). IEEE.. doi:10.1109/ROMAN.2015.7333613

*[110] Haring, K. S., Silvera-Tawil, D., Takahashi, T., Watanabe, K., & Velonaki, M. (2016, February). How people perceive different robot types: A direct comparison of an android, humanoid, and non-biomimetic robot. In *2016 8th International Conference on Knowledge and Smart Technology (KST)* (pp. 265-270). IEEE. doi:10.1109/KST.2016.7440504

Harmo, P., Taipalus, T., Knuuttila, J., Vallet, J., & Halme, A. (2005). Needs and solutions-home automation and service robots for the elderly and disabled. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3201-3206). IEEE. doi: 10.1109/IROS.2005.1545387

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological Bulletin, 135*(4), 555. doi: 10.1037/a0015701

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics, 2*(4), 361-375. doi: 10.1007/s12369-010-0068-5

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M., & Sagerer, G. (2009). Understanding social robots. In *2009 Second International Conferences on Advances in Computer-Human Interactions* (pp. 169-174). IEEE. doi: 10.1109/ACHI.2009.51

*[19] Herse, S., Vitale, J., Tonkin, M., Ebrahimian, D., Ojha, S., Johnston, B., & Williams, M. A. (2018). Do you trust me, blindly? Factors influencing trust towards a robot recommender system. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 7-14). IEEE. doi: 10.1109/ROMAN.2018.8525581

Hewstone, M., & Swart, H. (2011). Fifty-odd years of inter-group contact: From hypothesis to integrated theory. *British Journal of Social Psychology*, *50*(3), 374-386. doi: 10.1111/j.2044-8309.2011.02047.x

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*(2), 141-154. doi: 10.1016/S0022-1031(77)80007-3

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in human behavior, 49*, 245-250. doi: 10.1016/j.chb.2015.02.026

Hiroshi Ishiguro Laboratories, A. T. R. I. I. A. HIL. from http://www.geminoid.jp/

*[42] Hosseini, S. M. F., Lettinga, D., Vasey, E., Zheng, Z., Jeon, M., Park, C. H., & Howard, A. M. (2017). Both “look and feel” matter: Essential factors for robotic companionship. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 150-155). IEEE. doi: 10.1109/ROMAN.2017.8172294

IEEE (2010) History of ROMAN conferences [1992-2011], In *19th International Symposium in Robot and Human Interactive Communication, Viareggio, 2010*, 1. doi: 10.1109/ROMAN.2010.5598745

IEEE-RAS International Conference on Humanoid Robots (Humanoids). (2020). Retrieved January 25, 2020, from https://www.ieee-ras.org/conferences-workshops/fully-sponsored/humanoids

Im, I., Hong, S., & Kang, M. S. (2011). An international comparison of technology adoption: Testing the UTAUT model. *Information & management, 48*(1), 1-8. doi: 10.1016/j.im.2010.09.001

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., & Nakatsu, R. (2001). Robovie: an interactive humanoid robot. *Industrial robot: An international journal, 28*(6), 498-503. doi: 10.1108/01439910110410051

*[111] Iwamura, Y., Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2011). Do elderly people prefer a conversational humanoid as a shopping assistant partner in supermarkets?. In *Proceedings of the 6th international conference on Human-robot interaction*, 449-456. doi: 10.1145/1957656.1957816

Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M.-T. (2014). Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human-Computer Interaction, 30*(5), 369-393. doi: 10.1080/10447318.2013.873278

Kassebaum, U. B. (2004). *Interpersonelles Vertrauen: Entwicklung eines Inventars zur Erfassung spezifischer Aspekte des Konstrukts*. [Doctoral thesis, Universität Hamburg]. Open. http://www.sub.uni-hamburg.de/opus/volltexte/2004/2125

Katz, J. E., & Halpern, D. (2014). Attitudes towards robots suitability for various jobs as affected robot appearance. *Behaviour & Information Technology, 33*(9), 941-953. doi: 10.1080/0144929X.2013.783115

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871– 888. doi: 10.1037/0022-3514.78.5.871

Kegan, D. L., & Rubenstein, A. H. (1973). Trust, effectiveness, and organizational development: A field study in R & D. *The journal of applied behavioral science, 9*(4), 498-513. doi: 10.1177/002188637300900409

Kessler, T. T., Larios, C., Walker, T., Yerdon, V., & Hancock, P. (2017). A Comparison of Trust Measures in Human–Robot Interaction Scenarios. In *Advances in Human Factors in Robots and Unmanned Systems* (pp. 353-364): Springer. doi: 10.1007/978-3-319-41959-6_29

Khan, P. H., Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. Developmental Psychology, 48(2), 303–314. doi: 10.1037/a0027033

*[112] Khan, S., & Germak, C. (2018). Reframing HRI design opportunities for social robots: Lessons learnt from a service robotics case study approach using UX for HRI. *Future Internet, 10*(10), 101. doi:103390/fi10100101

*[113] Kim, K. J., Park, E., & Sundar, S. S. (2013). Caregiving role in human–robot interaction: A study of the mediating effects of perceived benefit and social presence. *Computers in Human Behavior, 29*(4), 1799-1806. doi:10.1016/j.chb.2013.02.009

*[47] Kim, M. G., Lee, J., Aichi, Y., Morishita, H., & Makino, M. (2016). Effectiveness of robot exhibition through visitors experience: A case study of Nagoya Science Hiroba exhibition in Japan. In *2016 International Symposium on Micro-NanoMechatronics and Human Science (MHS)* (pp. 1-5). IEEE.

Klamer, T., & Allouch, S. B. (2010). Acceptance and use of a social robot by elderly users in a domestic environment. In *2010 4th International Conference on Pervasive*

Computing Technologies for Healthcare (pp. 1-8). IEEE. doi: 10.4108/ICST.PERVASIVEHEALTH2010.8892

Klowait, N. (2018). The quest for appropriate models of human-likeness: anthropomorphism in media equation research. *Ai & Society, 33*(4), 527-536. doi: 10.1007/s00146-017-0746-z

*[114] Kraus, M., Kraus, J., Baumann, M., & Minker, W. (2018). Effects of Gender Stereotypes on Trust and Likability in Spoken Human-Robot Interaction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 112-118.

Kriz, S., Ferro, T. D., Damera, P., & Porter, J. R. (2010). Fictional robots as a data source in HRI research: Exploring the link between science fiction and interactional expectations. In *19th International Symposium in Robot and Human Interactive Communication* (pp. 458-463). IEEE. doi: 10.1109/ROMAN.2010.5598620

*[64] Kuchenbrandt, D., & Eyssel, F. (2012). The mental simulation of a human-robot interaction: Positive effects on attitudes and anxiety toward robots. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication* (pp. 463-468). IEEE. doi: 10.1109/ROMAN.2012.6343795

*[115] Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R. M. Q., & MacDonald, B. A. (2009). Age and gender factors in user acceptance of healthcare robots. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 214-219). IEEE. doi: 10.1109/ROMAN.2009.5326292

LaCour, M. J., & Green, D. P. (2014). When contact changes minds: An experiment on transmission of support for gay equality. *Science, 346*(6215), 1366-1369. doi: 10.1126/science.1256151

Lallée, S., Vouloutsi, V., Munoz, M. B., Grechuta, K., Llobet, J. Y. P., Sarda, M., & Verschure, P. F. (2015). Towards the synthetic self: making others perceive me as an other. *Paladyn, Journal of Behavioral Robotics, 6*(1). doi: 10.1515/pjbr-2015-0010

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: a survey. *International Journal of Social Robotics, 5*(2), 291-308. doi: 10.1007/s12369-013-0178-y

Lemmer, G., & Wagner, U. (2015). Can we really reduce ethnic prejudice outside the lab? A meta-analysis of direct and indirect contact interventions. *European Journal of Social Psychology, 45*(2), 152-168. doi: 10.1002/ejsp.2079

Lemon, N. (1973). *Attitudes and their measurement*. John Wiley & Sons.

*[20] Li, D., Rau, P. P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics, 2*(2), 175-186. doi: 10.1007/s12369-010-0056-9

*[116] Liang, Y., & Lee, S. A. (2016). Advancing the Strategic Messages Affecting Robot Trust Effect: The Dynamic of User-and Robot-Generated Content on Human–Robot Trust and Interaction Outcomes. *Cyberpsychology, Behavior, and Social Networking, 19*(9), 538-544. doi:10.1089/cyber.2016.0199

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22 140*, 55.

Løgstrup, K. E. (1997). *The ethical demand*. University of Notre Dame Press.

*[117] Louie, W. Y. G., McColl, D., & Nejat, G. (2014). Acceptance and attitudes toward a human-like socially assistive robot by older adults. *Assistive Technology*, *26*(3), 140-150. doi:101080/104004352013869703

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, *81*(5), 842. doi: 10.1037//0022-35I4.8I.5.842

*[118] Lugrin, B., Dippold, J., & Bergmann, K. (2018). Social Robots as a Means of Integration? an Explorative Acceptance Study considering Gender and Non-verbal Behaviour. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 2026-2032). IEEE. doi:10.1109/IROS.2018.8593818

Mori, M. (1970). The uncanny valley. *Energy*, *7*(4), 33-35.

*[119] MacDorman, K. F., Vasudevan, S. K., & Ho, C.-C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society*, *23*(4), 485-510. doi: 10.1007/s00146-008-0181-2

Mallett, R. K., & Wilson, T. D. (2010). Increasing positive intergroup contact. *Journal of Experimental Social Psychology*, *46*(2), 382-387. doi: 10.1016/j.jesp.2009.11.006

Marous, J. (2015, July 7). *Robots and AI Invade Banking*. The Financial Brand. https://thefinancialbrand.com/52735/robots-artificial-intelligence-ai-banking/

*[120] Matsui, T., & Yamada, S. (2018, August). Robot's Impression of Appearance and Their Trustworthy and Emotion Richness. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 88-93). IEEE. doi: 10.1109/ROMAN.2018.8525796

Matsusaka, Y. (2008). History and current researches on building a human interface for humanoid robots. In *Modeling communication with robots and virtual humans* (pp. 109-124). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-79037-2_6

*[121] Mavridis, N., Katsaiti, M. S., Naef, S., Falasi, A., Nuaimi, A., Araifi, H., & Kitbi, A. (2012). Opinions and attitudes toward humanoid robots in the Middle East. *AI & society, 27*(4), 517-534. doi: 101007/s00146-011-0370-2

May, D. C., Holler, K. J., Bethel, C. L., Strawderman, L., Carruth, D. W., & Usher, J. M. (2017). Survey of factors for the prediction of human comfort with a non-anthropomorphic robot in public spaces. *International Journal of Social Robotics, 9*(2), 165-180. doi: 10.1007/s12369-016-0390-7

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review, 20*(3), 709-734. doi: 10.5465/amr.1995.9508080335

McGuire, W. J. (1969). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd ed., Vol. 3, pp. 136-314). Reading, MA: Addison-Wesley

McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 233-346). New York: Random House

McKitterick, W. (2016). *The global smartphone report: The forces behind the global deceleration in smartphone sales.* Business Insider India. https://www.businessinsider.in/THE-GLOBAL-SMARTPHONE-REPORT-The-forces-behind-the-global-deceleration-in-smartphone-sales/articleshow/52018367.cms. Accessed 8 October 2019

Mewes, D., & Heloir, A. (2009). *The Uncanny Valley*.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of abnormal and social psychology, 67*(4), 371. doi: 10.1037/h0040525

Milgram, S., & Van den Haag, E. (1978). *Obedience to authority*. London: Tavistock.

*[122] Miller, K., Self, W., McNaughton, C., Mion, L., & Storrow, A. (2012). Attitudes Toward Health Care Robot Assistants In The ED: A Survey Of ED Patients And Visitors: 211. *Academic Emergency Medicine*, 19.

Milliman, R. E. (1982). Using background music to affect the behavior of supermarket shoppers. *Journal of Marketing, 46*(3), 86-91. doi: 10.1177/002224298204600313

*[123] Mirnig, N., Stadler, S., Stollnberger, G., Giuliani, M., & Tscheligi, M. (2016, August). Robot humor: How self-irony and Schadenfreude influence people's rating of robot likability. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 166-171). IEEE. doi: 10.1109/ROMAN.2016.7745106

Miskam, M. A., Hamid, M. A. C., Yussof, H., Shamsuddin, S., Malik, N. A., & Basir, S. N. (2013). Study on social interaction between children with autism and humanoid robot NAO. In *Applied mechanics and materials* (Vol. 393, pp. 573-578). Trans Tech Publications. doi: 10.4028/www.scientific.net/AMM.393.573

Misztal, B. (2013). *Trust in modern societies: The search for the bases of social order*. John Wiley & Sons.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*(3), 455-469 . doi: 10.1037/0096-3445.132.3.455

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J. J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning, 1*(209-0015), 13. doi: 10.2316/Journal.209.2013.1.209-0015

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues, 56*(1), 81-103. doi: 10.1111/0022-4537.00153

Nass, C., & Yen, C. (2010). *The man who lied to his laptop: What we can learn about ourselves from our machines*. Penguin.

Nations, U. (2007). World Population Prospects: The 2006 Revision, Highlights.

New Atlas. (2015, April 29) *NAO lends a robotic hand in banking customer service*. https://newatlas.com/japan-nao-robot-bank-customer-service/37053/

Nilsson, N. J. (1984). *Shakey the robot*. SRI INTERNATIONAL MENLO PARK CA.

*[61] Nomura, T. (2014). Influences of experiences of robots into negative attitudes toward robots. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 460-464). IEEE. doi: 10.1109/ROMAN.2014.6926295

*[124] Nomura, T. (2017). Cultural differences in social acceptance of robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 534-538). IEEE. doi:10.1109/ROMAN.2017.8172354

*[10] Nomura, T., Kanda, T., & Suzuki, T. (2006a). Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *Ai & Society*, *20*(2), 138-150. doi: 10.1007/s00146-005-0012-7

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2004). Psychology in human-robot communication: An attempt through investigation of negative attitudes and anxiety toward robots. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)* (pp. 35-40). IEEE. doi: 10.1109/ROMAN.2004.1374726

*[24] Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of Human Behavior in Human--Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. *IEEE transactions on robotics*, *24*(2), 442-451. doi: 10.1109/TRO.2007.914004

Nomura, T., Kanda, T., Yamada, S., & Suzuki, T. (2011). Exploring influences of robot anxiety into HRI. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 213-214). IEEE. doi: 10.1145/1957656.1957737

*[26] Nomura, T., Shintani, T., Fujii, K., & Hokabe, K. (2007). Experimental investigation of relationships between anxiety, negative attitudes, and allowable distance of robots. In *Proceedings of the 2nd IASTED international conference on human computer interaction, Chamonix, France. ACTA Press*, 13-18.

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006b). Measurement of negative attitudes toward robots. *Interaction Studies*, *7*(3), 437-454. doi: 10.1075/is.7.3.14nom

*[25] Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006c). Measurement of anxiety toward robots. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 372-377). IEEE. doi: 10.1109/ROMAN.2006.314462

Nørskov, M. (2017). Technological dangers and the potential of human–robot interaction: a philosophical investigation of fundamental epistemological mechanisms of discrimination. In M. Nørskov (Ed.), *Social Robots* (pp. 99-121). Routledge. doi: 10.4324/9781315563084

Nosek, B. A., & Lindsay, S. (2018, February 28). *Preregistration Becoming the Norm in Psychological Science*. https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science

Ortiz, M., & Harwood, J. (2007). A social cognitive theory approach to the effects of mediated intergroup contact on intergroup attitudes. *Journal of Broadcasting & Electronic Media*, *51*(4), 615-631. doi: 10.1080/08838150701626487

Ostrom, T. M. (1969). The relationship between the affective, behavioral, and cognitive components of attitude. *Journal of experimental social psychology*, *5*(1), 12-30. doi: 10.1016/0022-1031(69)90003-1

Pandey, A. K., & Gelin, R. (2018). A mass-produced sociable humanoid robot: Pepper: the first machine of its kind. *IEEE Robotics & Automation Magazine, 25*(3), 40-48. doi: 10.1109/MRA.2018.2833157

*[73] Park, E., & Del Pobil, A. P. (2012). An Acceptance Model for Service Robots in Global Markets. *International Journal of Humanoid Robotics, 9*(04), 1250026. doi: 10.1142/S0219843612500260

*[125] Park, E., & Kwon, S. J. (2016). The adoption of teaching assistant robots: a technology acceptance model approach. *Program: electronic library and information systems, 50*(4), 354-366. doi:10.1108/PROG-02-2016-0017

*[126] Park, E., & Lee, J. (2014). I am a warm robot: the effects of temperature in physical human–robot interaction. *Robotica, 32*(1), 133-142. doi:101017/S026357471300074X

Park, E., Jenkins, Q., & Jiang, X. (2008). Measuring trust of human operators in new generation rescue robots. In *JFPS International Symposium on Fluid power* (Vol. 2008, No. 7-2, pp. 489-492). The Japan Fluid Power System Society. doi: 10.5739/isfp.2008.489

Penn, D. (2016). True lies and true implicit: How priming reveals the hidden truth. *International Journal of Market Research, 58*(2), 175-199. doi: 10.2501/IJMR-2016-014

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual review of psychology, 49*(1), 65-85.

Pettigrew, T. F., Christ, O., Wagner, U., & Stellmacher, J. (2007). Direct and indirect intergroup contact effects on prejudice: A normative interpretation. *International Journal of intercultural relations, 31*(4), 411-425. doi: 10.1016/j.ijintrel.2006.11.003

Petty, R. E., Tormala, Z. L., Brinol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: an exploration of the PAST model. *Journal of Personality and Social Psychology, 90*(1), 21-41. doi: 10.1037/0022-3514.90.1.21

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), 751–783. doi: 10.1037/0022-3514.90.5.751

Pieska, S., Luimula, M., Jauhiainen, J., & Spiz, V. (2013). Social service robots in wellness and restaurant applications. *Journal of Communication and Computer, 10*(1), 116-123.

*[127] Pino, M., Boulay, M., Jouen, F., & Rigaud, A. S. (2015). "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in aging neuroscience, 7*, 141. doi:103389/fnagi201500141

Plous, S. (1993). *The psychology of judgment and decision making*. Mcgraw-Hill Book Company.

Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European review of social psychology, 17*(1), 271-320. doi: 10.1080/10463280601055772

Prescott, T. J. (2017). Robots are not just tools. *Connection Science, 29*(2), 142-149. doi: 10.1080/09540091.2017.1279125

Qureshi, A. H., Nakamura, Y., Yoshikawa, Y., & Ishiguro, H. (2016). Robot gains social intelligence through multimodal deep reinforcement learning. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)* (pp. 745-751). IEEE. doi: 10.1109/HUMANOIDS.2016.7803357

*[71] Rantanen, T., Lehto, P., Vuorinen, P., & Coco, K. (2018). Attitudes towards care robots among Finnish home care personnel–a comparison of two approaches. *Scandinavian journal of caring sciences, 32*(2), 772-782. doi: 101111/scs12508

Ray, C., Mondada, F., & Siegwart, R. (2008). What do people expect from robots?. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3816-3821. doi: 10.1109/IROS.2008.4650714

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge university press.

*[68] Reich-Stiebert, N., & Eyssel, F. (2015). Learning with educational companion robots? Toward attitudes on education robots, predictors of attitudes, and application potentials for education robots. *International Journal of Social Robotics, 7*(5), 875-888. doi: 10.1007/s12369-015-0308-9

*[38] Reich-Stiebert, N., Eyssel, F., & Hohnemann, C. (2019). Involve the user! Changing attitudes toward robots by user participation in a robot prototyping process. *Computers in Human Behavior, 91*, 290-296. doi: 10.1016/j.chb.2018.09.041

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of personality and social psychology, 49*(1), 95.

Reuten, A., van Dam, M., & Naber, M. (2018). Pupillary responses to robotic and human emotions: the uncanny valley and media equation confirmed. *Frontiers in Psychology, 9*, 774. doi: 10.3389/fpsyg.2018.00774

Ring, P. S., & Van de Ven, A. H. (1994). Developmental processes of cooperative interorganizational relationships. *Academy of management review, 19*(1), 90-118. doi: 10.5465/amr.1994.9410122009

*[128] Robert, D., & van den Bergh, V. (2014). Children's Openness to Interacting with a Robot Scale (COIRS). In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 930-935). IEEE. doi: 10.1109/ROMAN.2014.6926372

Robins, B., Dautenhahn, K., & Dubowski, J. (2004). Investigating Autistic children's attitudes towards strangers with the theatrical robot-A new experimental paradigm in human-robot interaction studies. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)* (pp. 557-562). IEEE. doi: 10.1109/ROMAN.2004.1374821

Robot Advance. (2016). *Alpha 1E Robot Ubtech: programmable and customizable*. https://www.robot-advance.com/EN/art-ubtech-robot-alpha-1e-en-2671.htm

Rose, A. M. (1947). *Studies in Reduction of Prejudice*. American. Council on Race Relations.

Roy, N., Baltus, G., Fox, D., Gemperle, F., Goetz, J., Hirsch, T., & Thrun, S. (2000). Towards personal service robots for the elderly. In *Workshop on Interactive Robots and Entertainment (WIRE 2000)* (Vol. 25, p. 184).

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science, 17*(11), 954-958. doi: 10.1111/j.1467-9280.2006.01811.x

Šabanović, S. (2010). Robots in society, society in robots. *International Journal of Social Robotics, 2(4),* 439-450. doi: 10.1007/s12369-010-0066-7

Saenger, G. (1953). *The social psychology of prejudice.* New York: Harper.

Salem, M., Eyssel, F., Rohlfing, K., Kopp, S., & Joublin, F. (2013). To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics, 5*(3), 313-323. doi:10.1007/s12369-013-0196-9

*[129] Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 1-8). IEEE.

Savela, N., Turja, T., & Oksanen, A. (2018). Social acceptance of robots in different occupational fields: A systematic literature review. *International Journal of Social Robotics, 10*(4), 493-502. doi: 10.1007/s12369-017-0452-5

Schaefer, K. E. (2013). *The perception and measurement of human-robot trust.* [Doctoral dissertation, University of Central Florida]. Open. http://purl.fcla.edu/fcla/etd/CFE0004931

Schiappa, E., Gregg, P. B., & Hewes, D. E. (2005). The parasocial contact hypothesis. *Communication Monographs, 72*(1), 92-115. doi: 10.1080/0363775052000342544

Schutte, M. (2019). *Socially assistive robots in the elderly care: The attitudes of healthcare professionals towards the use of socially assistive robots.* [Master's thesis, University of Twente]. Open. http://purl.utwente.nl/essays/79233

Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: a psychological approach. *Universal access in the information society, 4*(2), 146-155. doi: 10.1007/s10209-005-0118-1

*[49] Serholt, S., Basedow, C. A., Barendregt, W., & Obaid, M. (2014). Comparing a humanoid tutor to a human tutor delivering an instructional task to children. In *2014 IEEE-RAS International Conference on Humanoid Robots* (pp. 1134-1141). IEEE.

Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A., & Zahari, N. I. (2012a). Humanoid robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills. *Procedia Engineering, 41*, 1533-1538. doi: 10.1016/j.proeng.2012.07.346

Shamsuddin, S., Yussof, H., Ismail, L., Hanapiah, F. A., Mohamed, S., Piah, H. A., & Zahari, N. I. (2012b). Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO. In *2012 IEEE 8th International Colloquium on Signal Processing and its Applications* (pp. 188-193). IEEE. doi: 10.1109/CSPA.2012.6194716

Sherif, M. (2010). *The Robbers Cave Experiment [electronic resource]: Intergroup Conflict and Cooperation. [Orig. pub. as Intergroup Conflict and Group Relations]*. Middletown: Wesleyan University Press.

*[59] Shin, D. H., & Choo, H. (2011). Modeling the acceptance of socially interactive robotics: Social presence in human–robot interaction. Interaction Studies, 12(3), 430-460. doi: 10.1075/is.12.3.04shi

Simpson, G. E., & Yinger, J.M. (1965). *Racial and cultural minorities.* (3rd ed.). New York: Harper & Row.

Skinner, A. L., & Cheadle, J. E. (2016). The "Obama effect"? Priming contemporary racial milestones increases implicit racial bias among Whites. *Social Cognition, 34*(6), 544-558. doi: 10.1521/soco.2016.34.6.544

Softbank Robotics (2019). *Pepper the humanoid and programmable robot*. https://www.softbankrobotics.com/emea/en/pepper

*[74] Spence, P. R., Edwards, A., & Edwards, C. (2018). Attitudes, prior interaction, and petitioner credibility predict support for considering the rights of robots. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 243-244. doi: 10.1145/3173386.3177071

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of personality and social psychology, 37*(10), 1660. doi: 10.1037/0022-3514.37.10.1660

*[130] Stadler, S., Weiss, A., & Tscheligi, M. (2014). I trained this robot: the impact of pre-experience and execution behavior on robot teachers. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 1030-1036). IEEE. doi:10.1109/ROMAN.2014.6926388

Stafford, R. Q., Broadbent, E., Jayawardena, C., Unger, U., Kuo, I. H., Igic, A., & MacDonald, B. A. (2010). Improved robot attitudes and emotions at a retirement home after meeting a robot. In *19th international symposium in robot and human interactive communication* (pp. 82-87). IEEE. doi: 10.1109/ROMAN.2010.5598679

*[66] Stanton, C. J., & Stevens, C. J. (2017). Don't stare at me: the impact of a humanoid robot's gaze upon trust during a cooperative human–robot visual task. *International Journal of Social Robotics, 9*(5), 745-753. doi: 10.1007/s12369-017-0422-y

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006, March). Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction* (pp. 33-40). doi: 10.1145/1121241.1121249

Steinmetz, H., Knappstein, M., Ajzen, I., Schmidt, P., & Kabst, R. (2016). How effective are behavior change interventions based on the theory of planned behavior?. *Zeitschrift für Psychologie, 224*(3), 216–233. doi: 10.1027/2151-2604/a000255

*[131] Stuck, R. E., & Rogers, W. A. (2018). Older adults' perceptions of supporting factors of trust in a robot care provider. *Journal of Robotics, 2018*. doi:10.1155/2018/6519713

*[132] Sundar, S. S., Waddell, T. F., & Jung, E. H. (2016). The Hollywood Robot Syndrome media effects on older adults' attitudes toward robots and adoption intentions. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 343-350). IEEE. doi:101109/HRI20167451771

Syrdal, D. S., Dautenhahn, K., Koay, K. L., & Ho, W. C. (2014). Views from within a narrative: Evaluating long-term human–robot interaction in a naturalistic environment using open-ended scenarios. *Cognitive computation, 6*(4), 741-759. doi: 10.1007/s12559-014-9284-x

Syrdal, D. S., Dautenhahn, K., Koay, K. L., & Walters, M. L. (2009). The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study. *Adaptive and Emergent Behaviour and Complex Systems*.

Taheri, A. R., Alemi, M., Meghdari, A., Pouretemad, H. R., & Holderread, S. L. (2015). Clinical application of humanoid robots in playing imitation games for autistic children in Iran. *Procedia-Social and Behavioral Sciences, 176*, 898-906. doi: 10.1016/j.sbspro.2015.01.556

Takeda, T., Hirata, Y., & Kosuge, K. (2007). Dance step estimation method based on HMM for dance partner robot. *IEEE Transactions on Industrial Electronics, 54*(2), 699-706. doi: 10.1109/TIE.2007.891642

Tanaka, F., & Kimura, T. (2009). The use of robots in early education: a scenario based on ethical consideration. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 558-560). IEEE. doi: 10.1109/ROMAN.2009.5326227

Tanaka, F., Isshiki, K., Takahashi, F., Uekusa, M., Sei, R., & Hayashi, K. (2015). Pepper learns together with children: Development of an educational application. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)* (pp. 270-275). IEEE. doi: 10.1109/HUMANOIDS.2015.7363546

Tapus, A., Peca, A., Aly, A., Pop, C., Jisa, L., Pintea, S., Rusu, A., & David, D. O. (2012). Children with autism social engagement in interaction with Nao, an imitative robot: A series of single case experiments. *Interaction studies, 13*(3), 315-347. doi: 10.1075/is.13.3.01tap

*[60] Tay, B., Jung, Y., & Park, T. (2014). When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior, 38*, 75-84. doi: 10.1016/j.chb.2014.05.014

The Guardian. (2015, February 4). *Japanese bank introduces robot workers to deal with customers in branches*. https://www.theguardian.com/world/2015/feb/04/japanese-bank-introduces-robot-workers-to-deal-with-customers-in-branches

Thrun, S. (2004). Toward a framework for human-robot interaction. *Human–Computer Interaction, 19*(1-2), 9-24.

Thurstone, L. L. (1928). Attitudes can be measured. *American journal of Sociology, 33*(4), 529-554.

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior, 27*(2), 741-749. doi: 10.1016/j.chb.2010.10.018

*[79] Torta, E., Oberzaucher, J., Werner, F., Cuijpers, R. H., & Juola, J. F. (2013). Attitudes towards socially assistive robots in intelligent homes: results from laboratory studies and field trials. *Journal of Human-Robot Interaction, 1*(2), 76-99. doi: 10.5898/JHRI.1.2.Torta

*[43] Torta, E., Werner, F., Johnson, D. O., Juola, J. F., Cuijpers, R. H., Bazzani, M., & Bregman, J. (2014). Evaluation of a small socially-assistive humanoid robot in intelligent homes for the care of the elderly. *Journal of Intelligent & Robotic Systems, 76*(1), 57-71. doi:101007/s10846-013-0019-0

Traeger, M. L., Sebo, S. S., Jung, M., Scassellati, B., & Christakis, N. A. (2020). Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proceedings of the National Academy of Sciences, 117*(12), 6370-6375. doi: 10.1073/pnas.1910402117

*[133] Tung, F. W. (2016). Child perception of humanoid robot appearance and behavior. *International Journal of Human-Computer Interaction, 32*(6), 493-502. doi: 10.1080/10447318.2016.1172808

Turing, A. (1950), Computing Machinery and Intelligence, *Mind*, *59*(236), pp. 433–460

Turkle, S. (1998). Cyborg babies and cy-dough-plasm: ideas about life in the culture of simulation. In: R. Davis-Floyd, & J. Dumit J. (Eds.), *Cyborg babies: from techno-sex to techno-tots*. (pp 317–329). Routledge.

Turkle, S. (2005). *The second self: Computers and the human spirit*. MIT Press, London.

Turkle, S. (2017, December 7). *Why these friendly robots can't be good friends to our kids*. The Washington Post. https://www.washingtonpost.com

Turkle, S., Helmreich, S. (2009). *Simulation and Its Discontents*. MIT Press, London.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS quarterly*, 115-139. doi: 10.2307/3250981

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly, 27*(3), 425–478. doi: 10.2307/30036540

Vercelli, A., Rainero, I., Ciferri, L., Boido, M., & Pirri, F. (2018). Robots in elderly care. *DigitCult-Scientific Journal on Digital Cultures*, *2*(2), 37-50. doi: 10.4399/97888255088954

*[135] Volante, W. G., Sosna, J., Kessler, T., Sanders, T., & Hancock, P. A. (2019). Social conformity effects on trust in simulation-based human-robot interaction. *Human factors*, *61*(5), 805-815. doi:10.1177/0018720818811190

Volante, W. G., Sanders, T., Dodge, D., Yerdon, V. A., & Hancock, P. A. (2016). Specifying influences that mediate trust in human-robot interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 1755-1759). Sage CA: Los Angeles, CA: SAGE Publications. doi: 10.1177/1541931213601402

Von Der Pütten, A. M., & Krämer, N. C. (2014). How design characteristics of robots determine evaluation and uncanny valley related responses. *Computers in Human Behavior*, *36*, 422-439. doi: 10.1016/j.chb.2014.03.066

Von der Pütten, A. M., & Krämer, N. C. (2015). Individuals' Evaluations of and Attitudes Towards Potentially Uncanny Robots. *International Journal of Social Robotics, 7*(5), 799-824. doi: 10.1007/s12369-015-0321-z

Von der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S. H. (2010). "It doesn't matter what you are!" explaining social effects of agents and avatars. *Computers in Human Behavior, 26*(6), 1641-1650. doi: 10.1016/j.chb.2010.06.012

Wahde, M. (2007). *Introduction to autonomous robots*. Department of Applied Mechanics, Chalmers University of Technology, Goteborg, Sweden.

Warwick, K., & Shah, H. (2016). Can machines think? A report on Turing test experiments at the Royal Society. *Journal of experimental & Theoretical artificial Intelligence, 28*(6), 989-1007. doi: 10.1080/0952813X.2015.1055826

*[136] Weiss, A., Bernhaupt, R., Tscheligi, M., Wollherr, D., Kuhnlenz, K., & Buss, M. (2008). A methodological variation for acceptance evaluation of human-robot interaction in public places. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 713-718). IEEE. doi: 10.1109/ROMAN.2008.4600751

*[137] Weiss, A., Bernhaupt, R., Tscheligi, M., & Yoshida, E. (2009). Addressing user experience and societal impact in a user study with a humanoid robot. In *AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction, SSAISB*, 150-157.

Weiss, A., & Bartneck, C. (2015). Meta analysis of the usage of the godspeed questionnaire series. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 381-388). IEEE. doi: 10.1109/ROMAN.2015.7333568

West, K., & Turner, R. (2014). Using extended contact to improve physiological responses and behavior toward people with schizophrenia. *Journal of Experimental Social Psychology*, *50*, 57-64. doi: 10.1016/j.jesp.2013.06.009

Williams Jr, R. M. (1947). The reduction of intergroup tensions: a survey of research on problems of ethnic, racial, and religious group relations. *Social Science Research Council Bulletin, 57*, xi, 153.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*(1), 101. doi: 10.1037//0033-295X.107.1.101

Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. Guilford Press.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*(5), 815-827. doi: 10.1037//0022-3614.81.5.815

WowWee (2020) *CHiP: The smart and lovable robot dog*. https://wowwee.com/chip

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology, 73*(1), 73-90. doi: 10.1037/0022-3514.73.1.73

*[63] Wu, Y. H., Wrobel, J., Cornuet, M., Kerhervé, H., Damnée, S., & Rigaud, A. S. (2014). Acceptance of an assistive robot in older adults: a mixed-method study of human–robot interaction over a 1-month period in the Living Lab setting. *Clinical interventions in aging*, *9*, 801. doi: 10.2147/CIA.S56435

*[65] Wullenkord, R., & Eyssel, F. (2014). Improving attitudes towards social robots using imagined contact. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 489-494). IEEE. doi: 10.1109/ROMAN.2014.6926300

*[72] Wullenkord, R., Fraune, M. R., Eyssel, F., & Šabanović, S. (2016). Getting in touch: How imagined, actual, and physical contact affect evaluations of robots. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 980-985). IEEE. doi: 10.1109/ROMAN.2016.7745228

Zajonc, R. B. (2017). The concepts of balance, congruity, and dissonance. In A. Pelinka (Ed.), *Attitude Change* (pp. 63-85). Routledge. doi: 10.4324/9781351315364

Zhao, D., & Yi, J. (2006). Robot planning with artificial potential field guided ant colony optimization algorithm. In *International Conference on Natural Computation* (pp. 222-231). Springer. doi: 10.1007/11881223_28

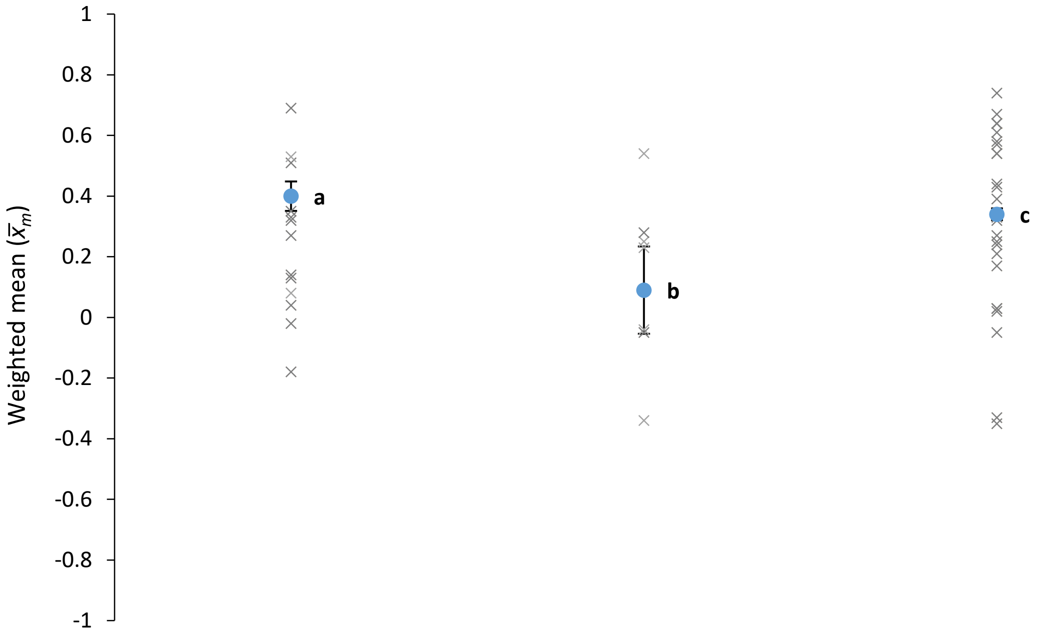
Zhou, S., Page-Gould, E., Aron, A., Moyer, A., & Hewstone, M. (2019). The extended contact hypothesis: A meta-analysis on 20 years of research. *Personality and Social Psychology Review*, *23*(2), 132-160. doi: 10.1177/1088868318762647

*[39] Zlotowski, J. A., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2015). Persistence of the uncanny valley: the influence of repeated interactions and a robot's attitude on its perception. *Frontiers in psychology*, *6*, 883. doi: 10.3389/fpsyg.2015.00883

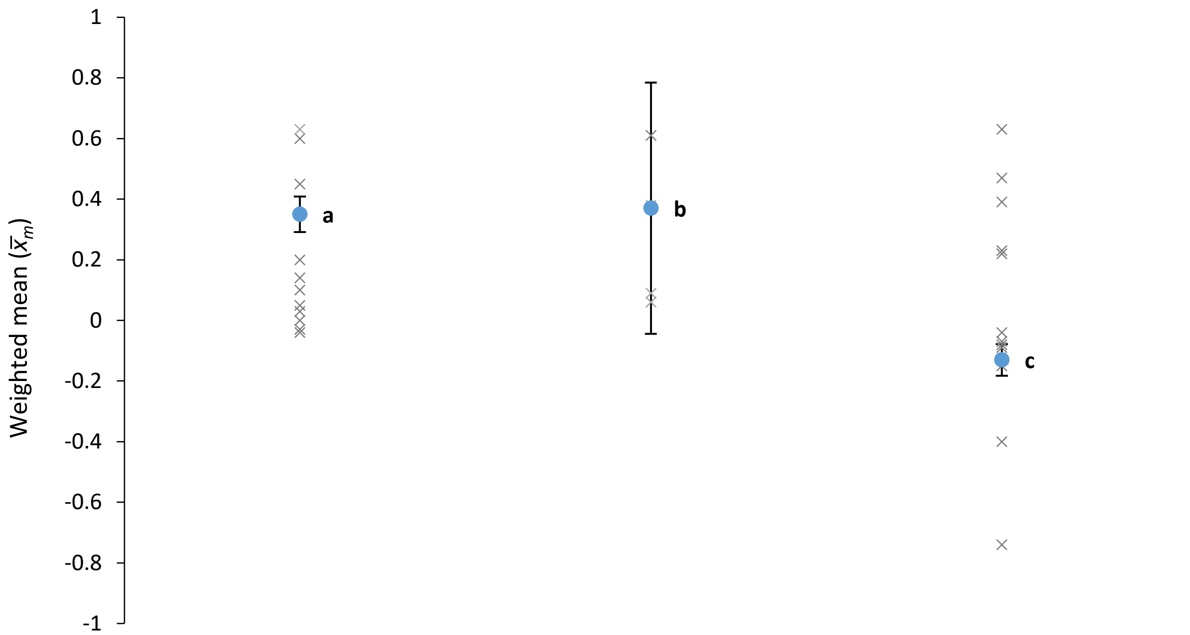
Appendix:

Supplementary Materials 1: Moderator graphs

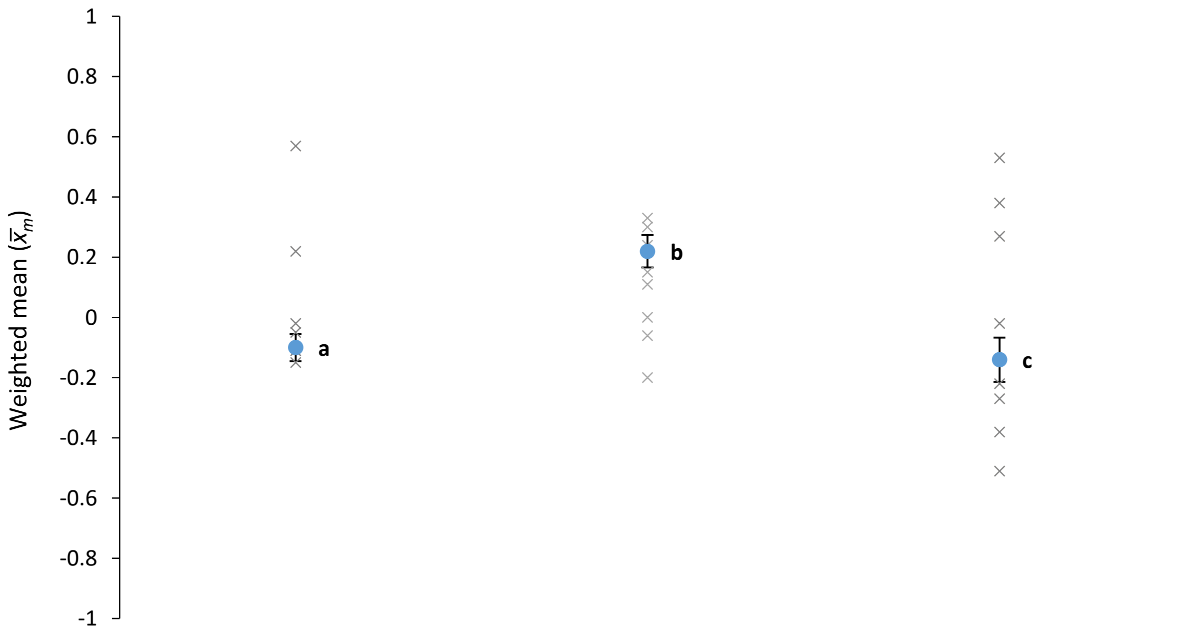
<u>Type of exposure to robots</u>



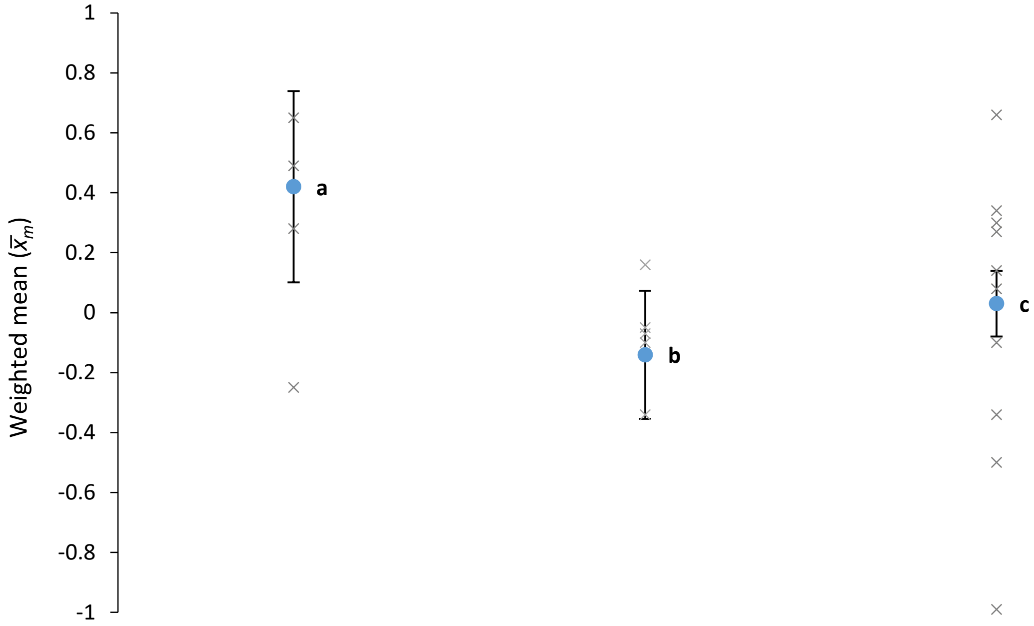
Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.



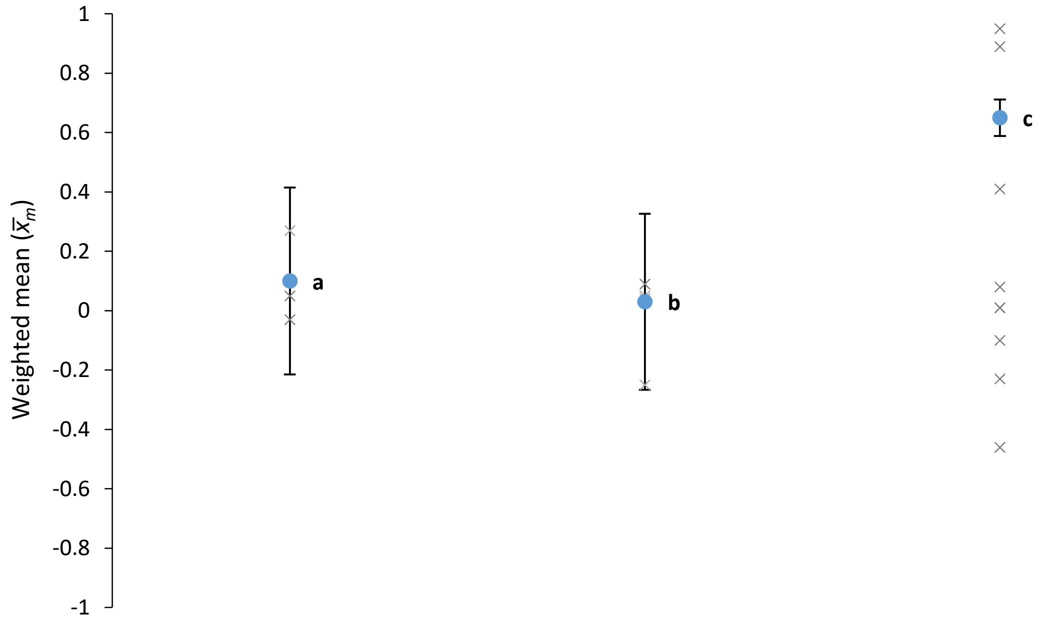
Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.



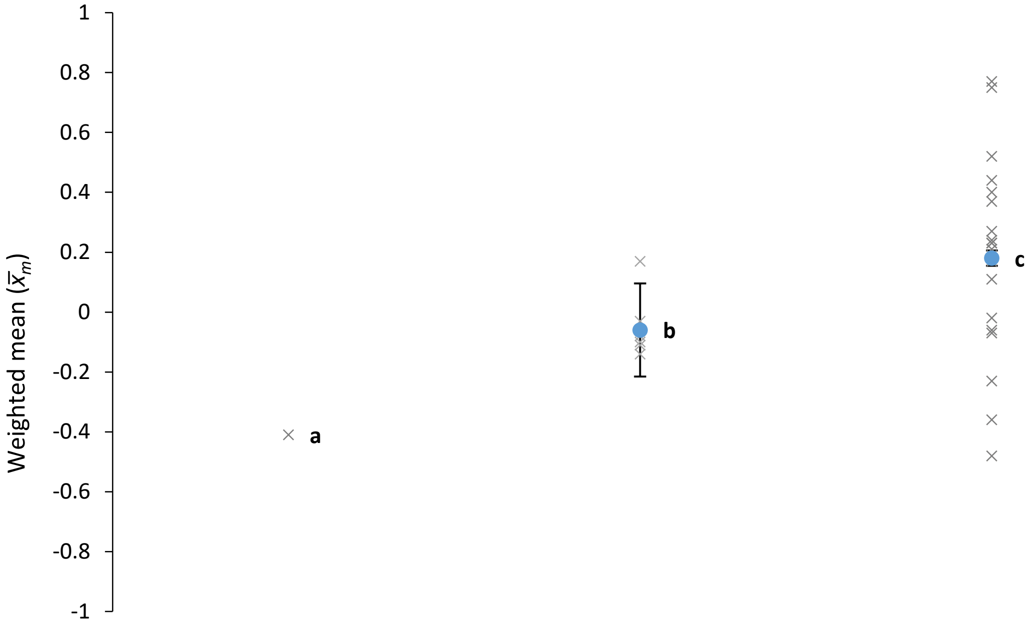
General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.



Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

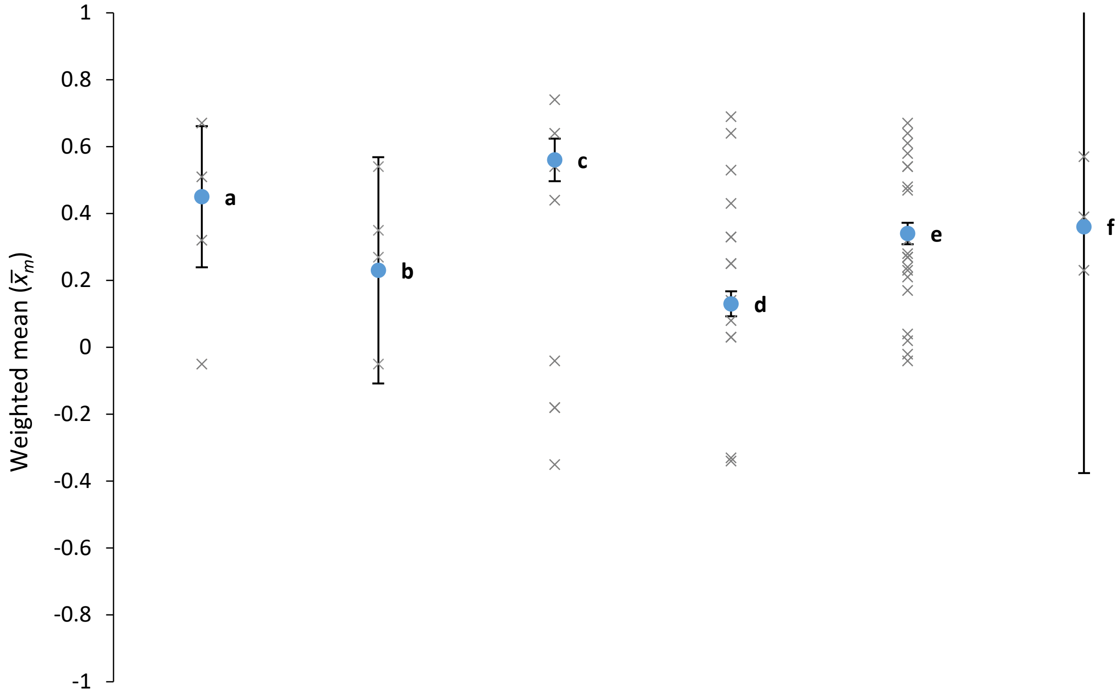


Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

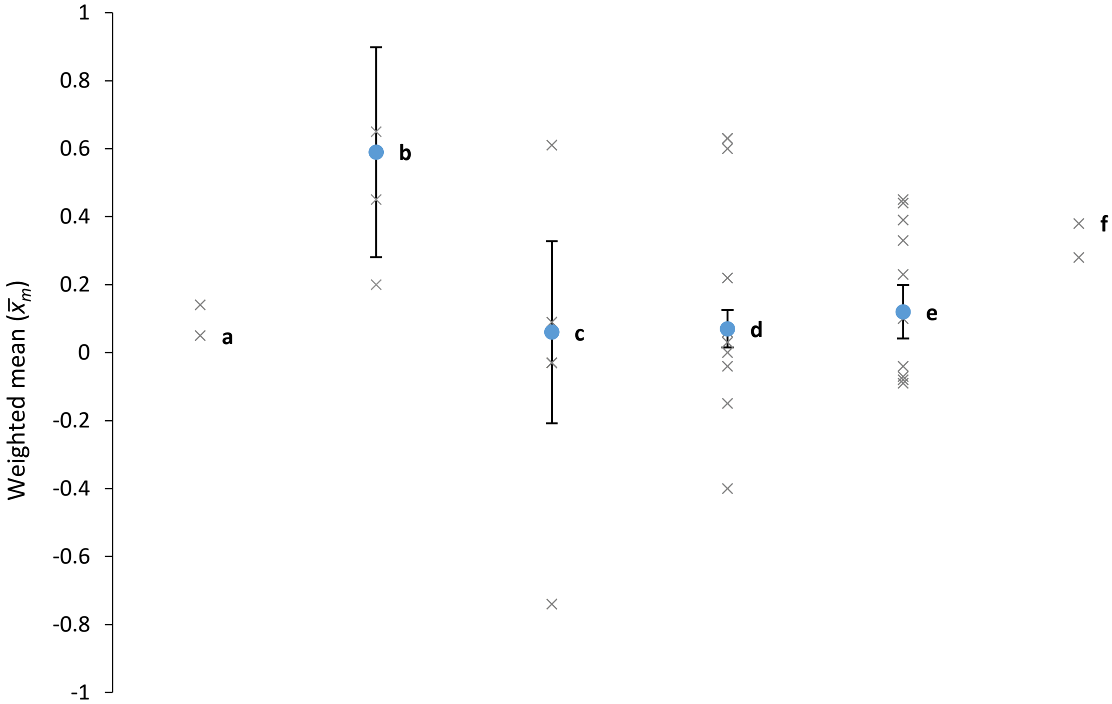


Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

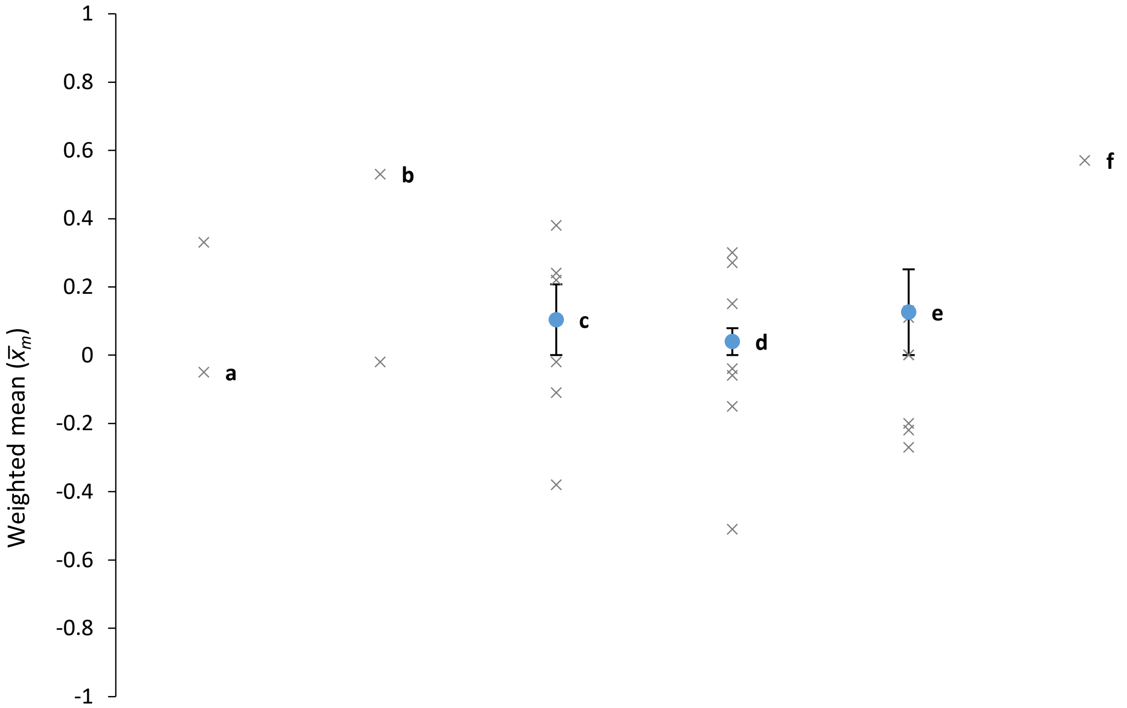
Domain of application



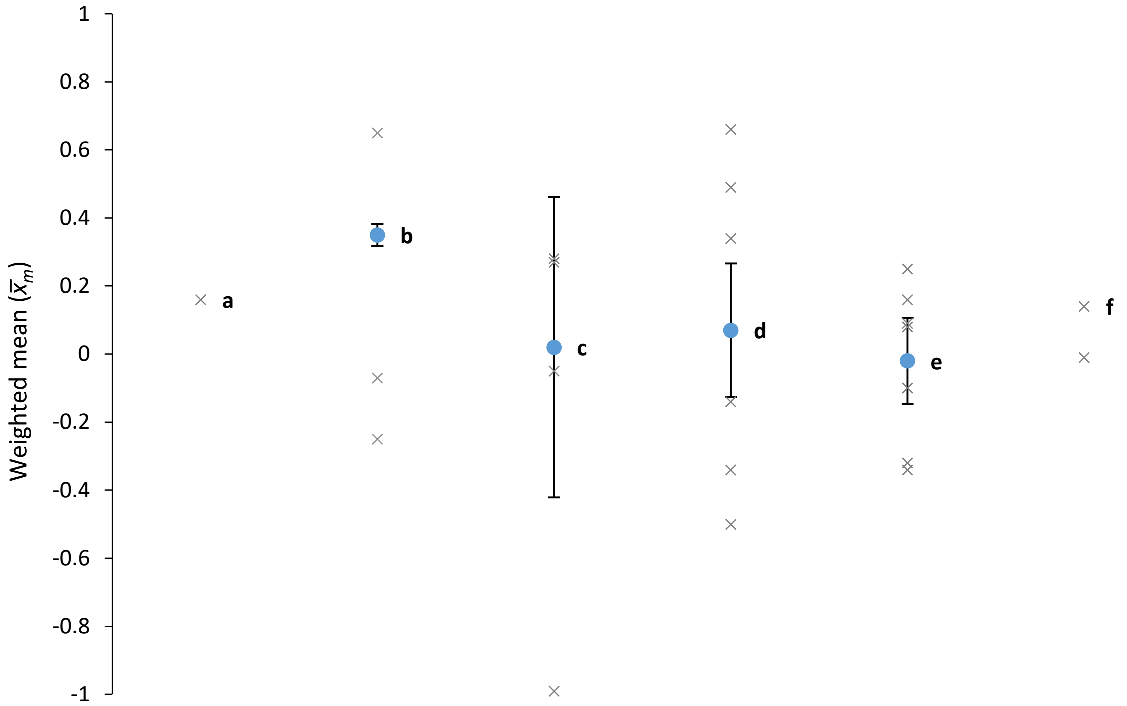
Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.



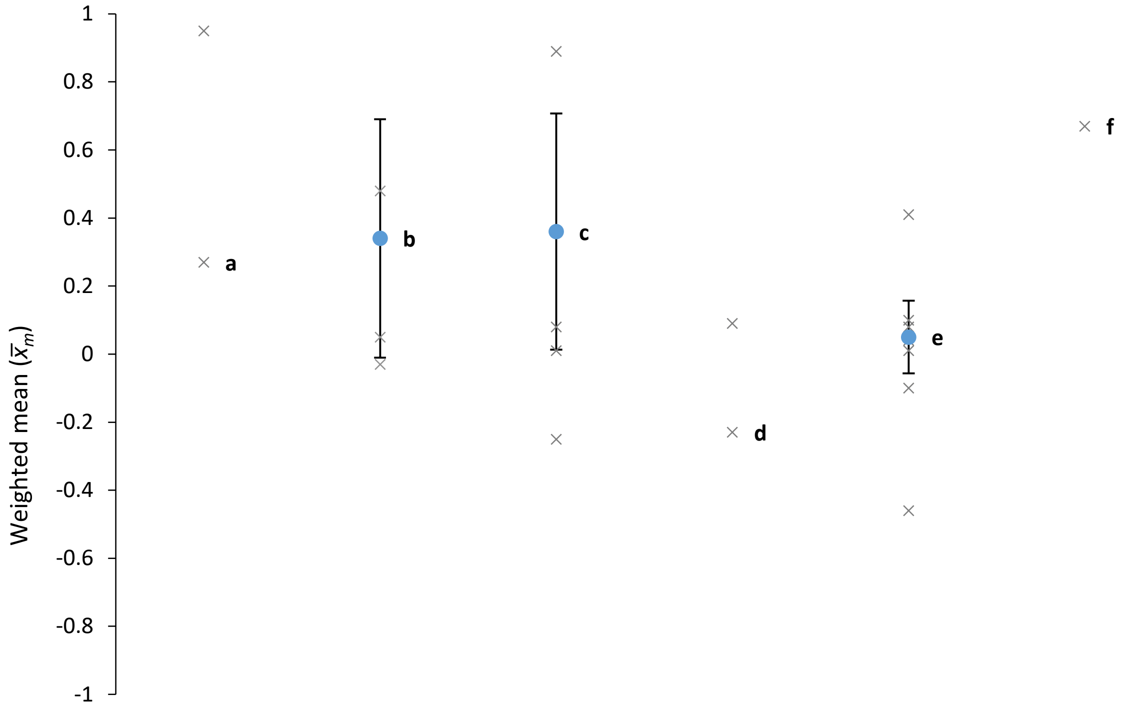
Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.



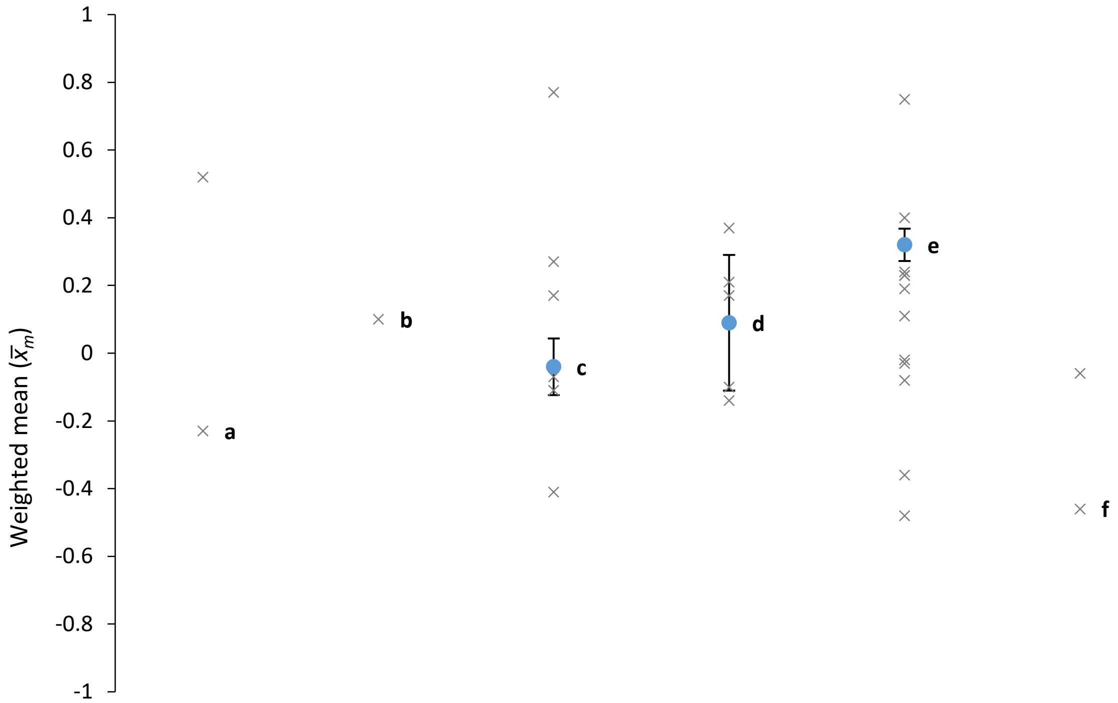
General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.



Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.

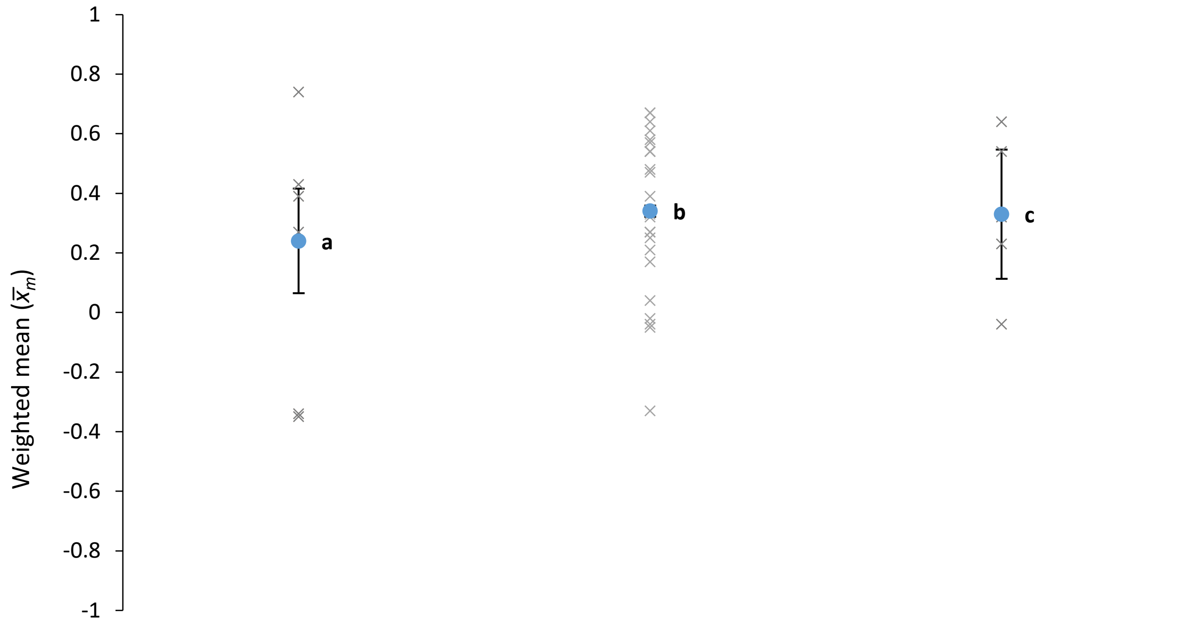


Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.

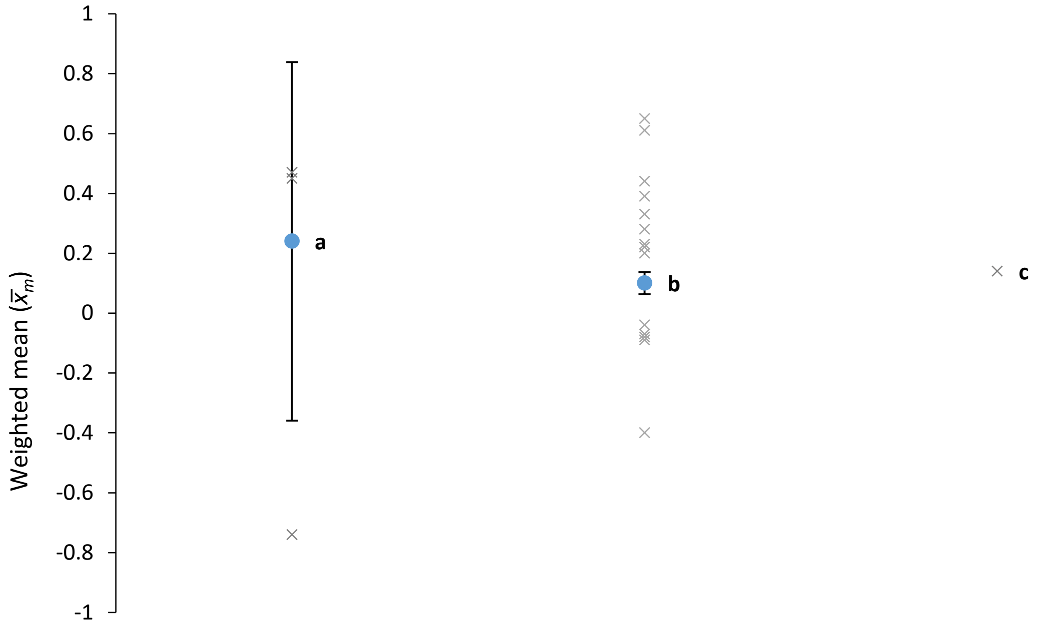


Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Companion robots & Domestic assistance group; b) Education group; c) Healthcare group; d) General application group; e) HRI group; and f) Pediatric care group.

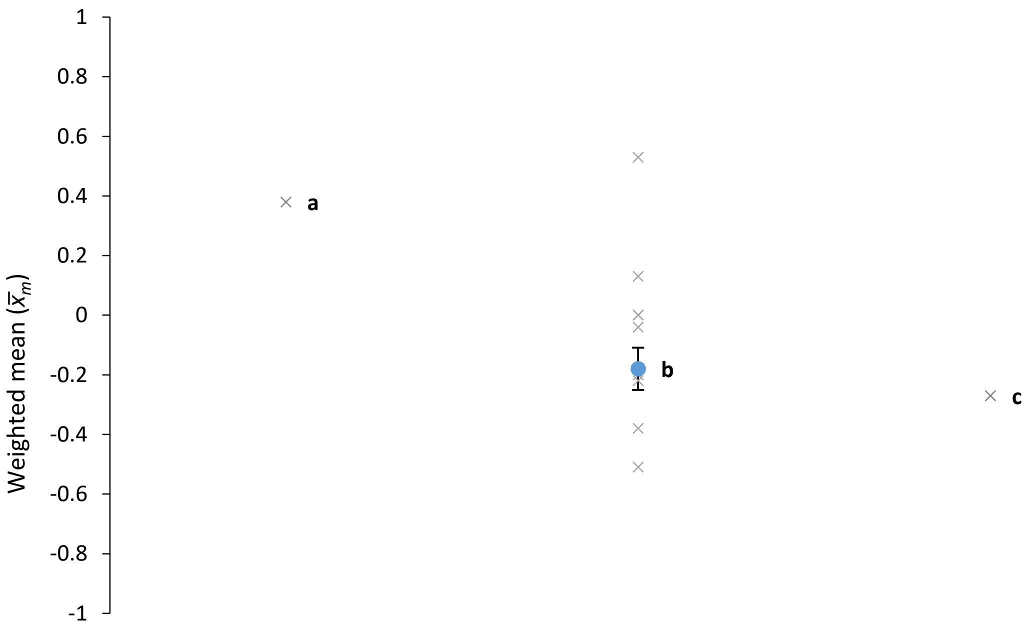
Design of robot



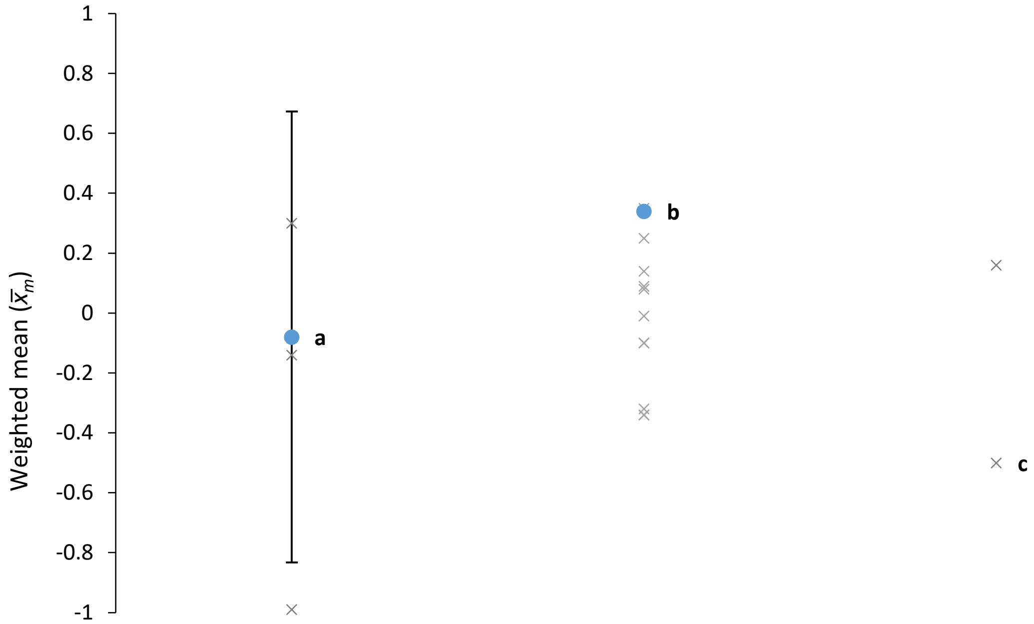
Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.



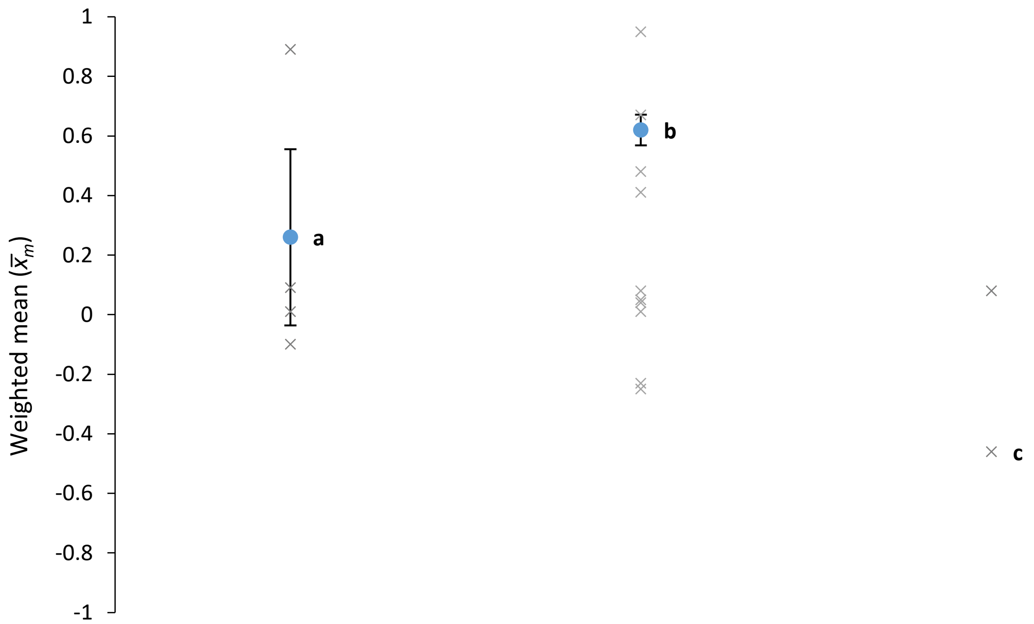
Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.



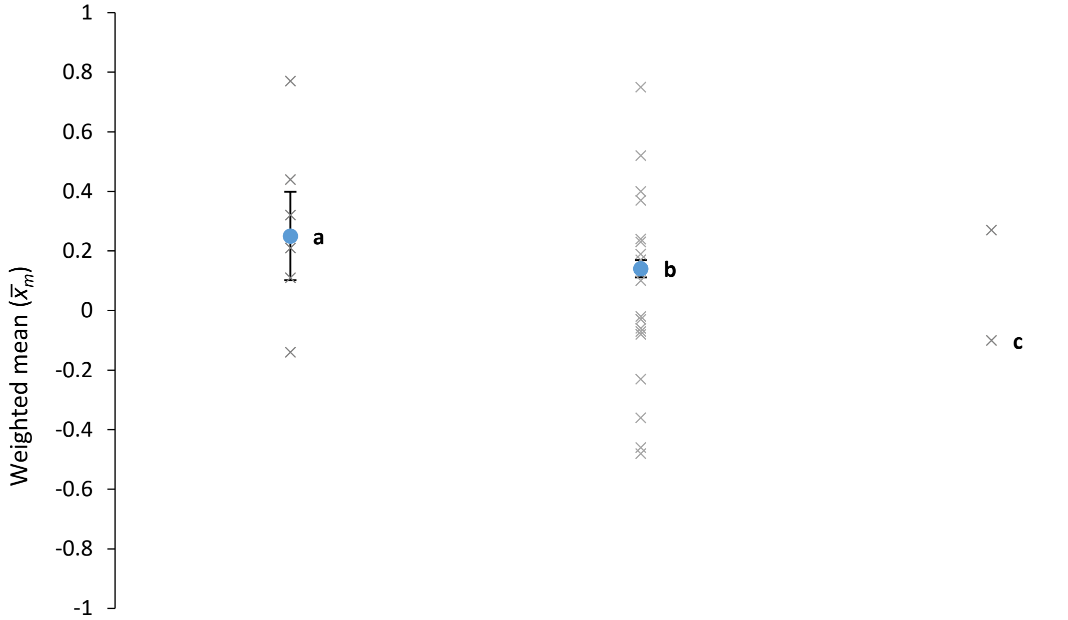
General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.



Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

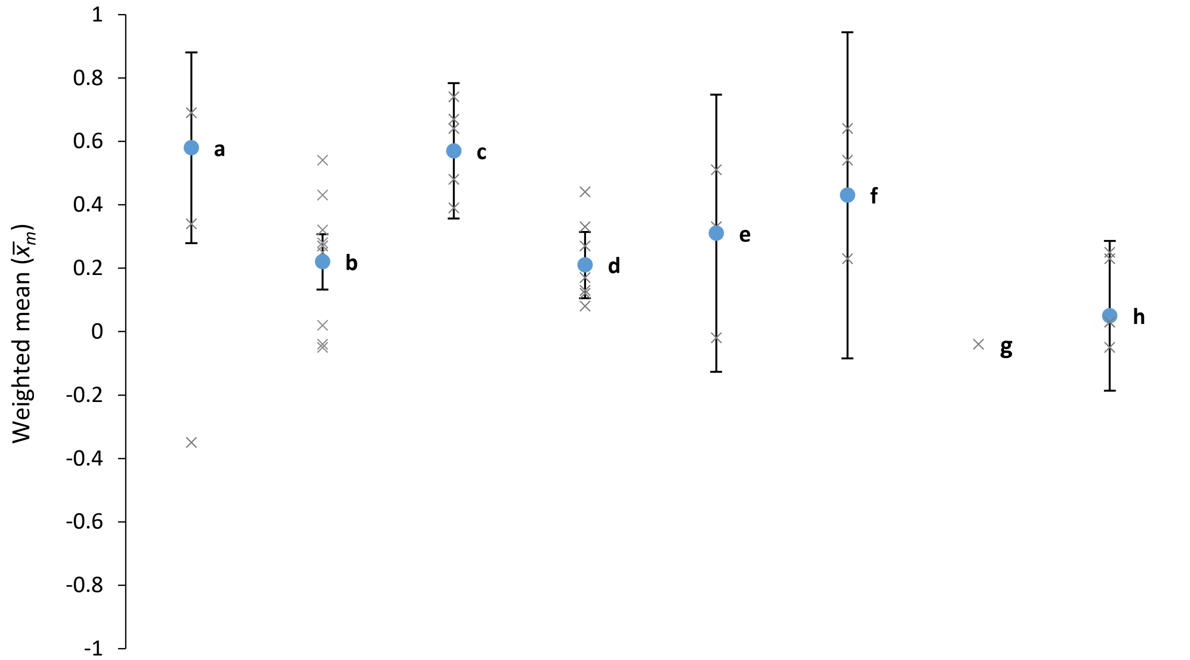


Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

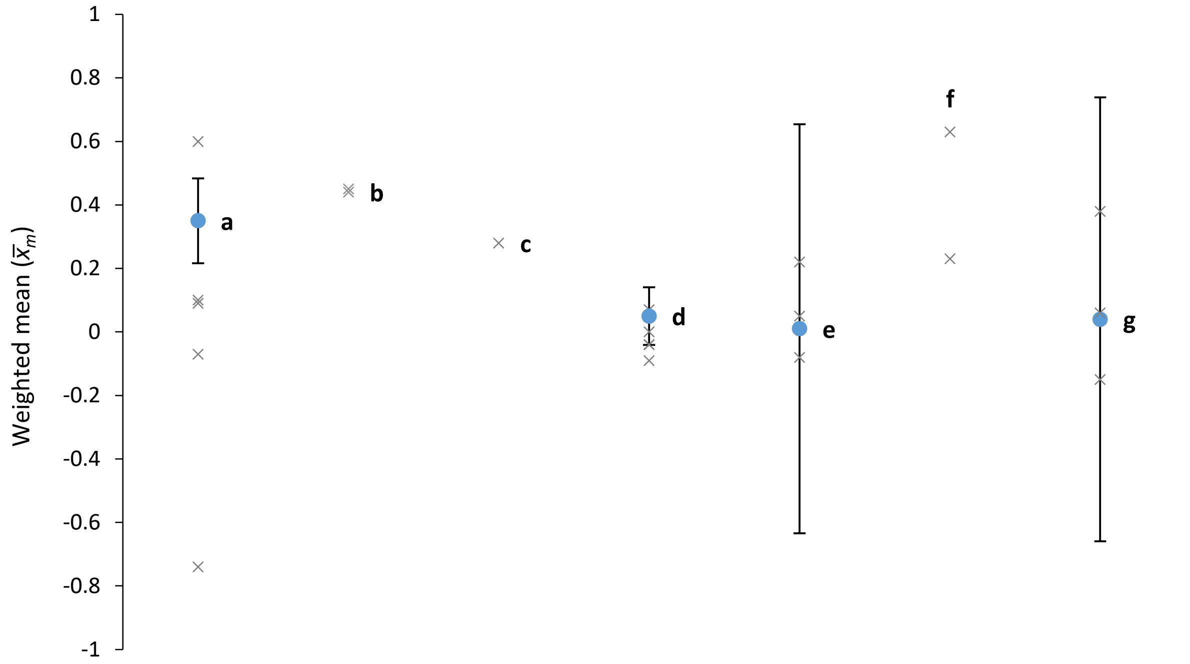


Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

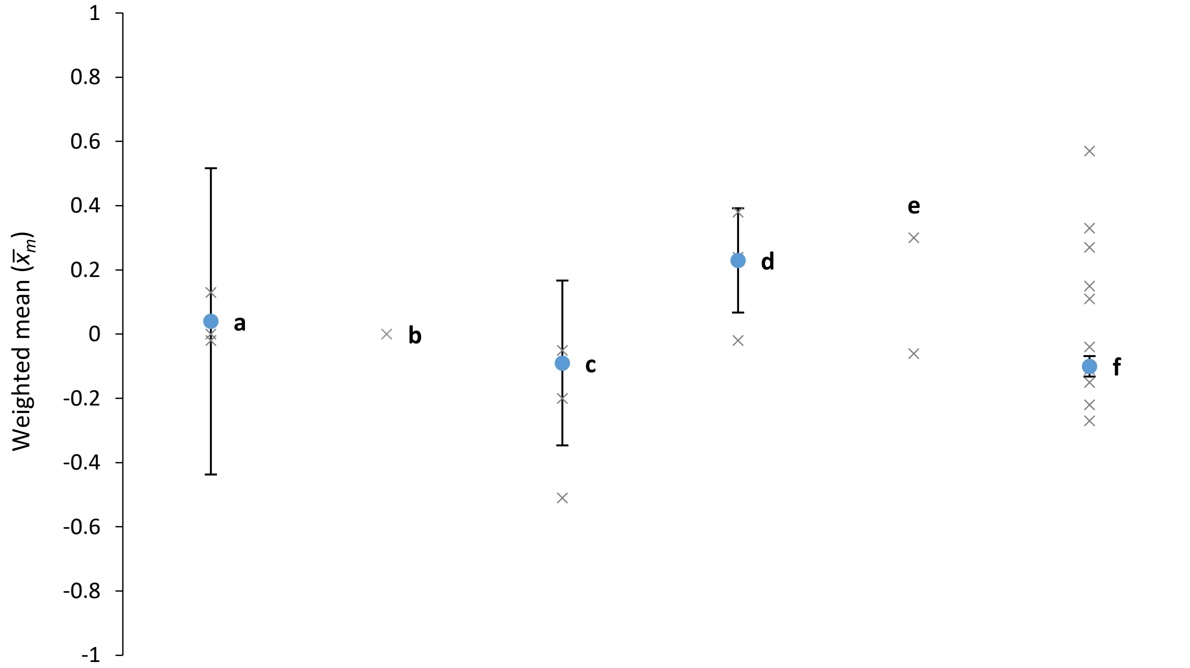
Country in which the research was conducted



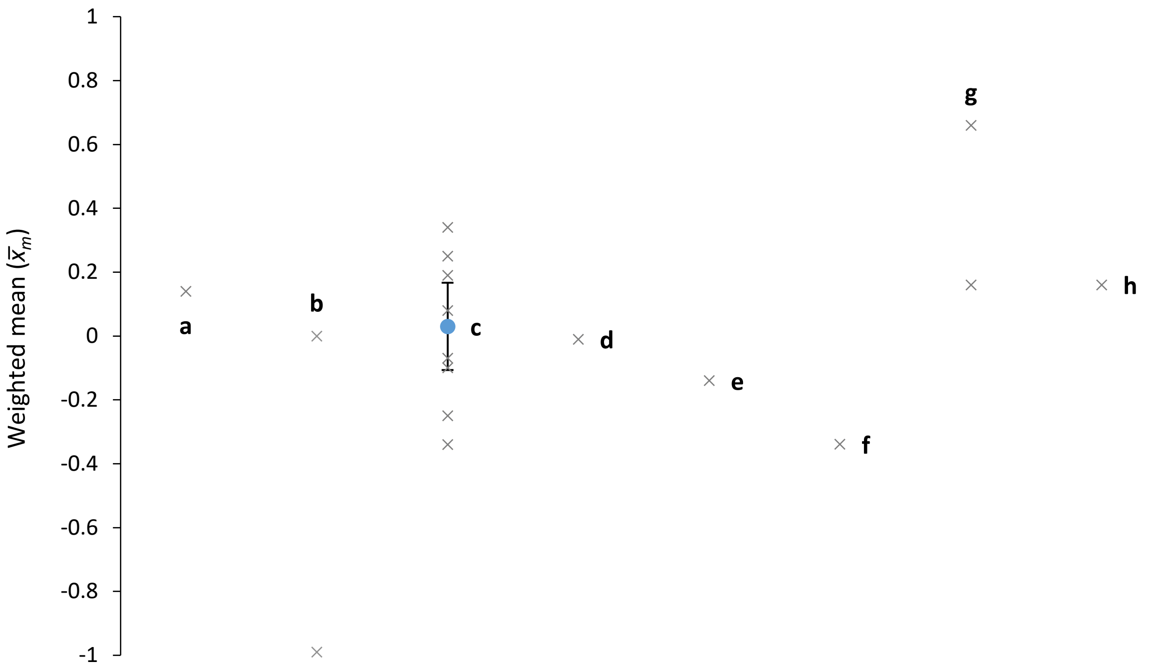
Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; e) Netherlands; f) South Korea; g) Taiwan; and h) USA.



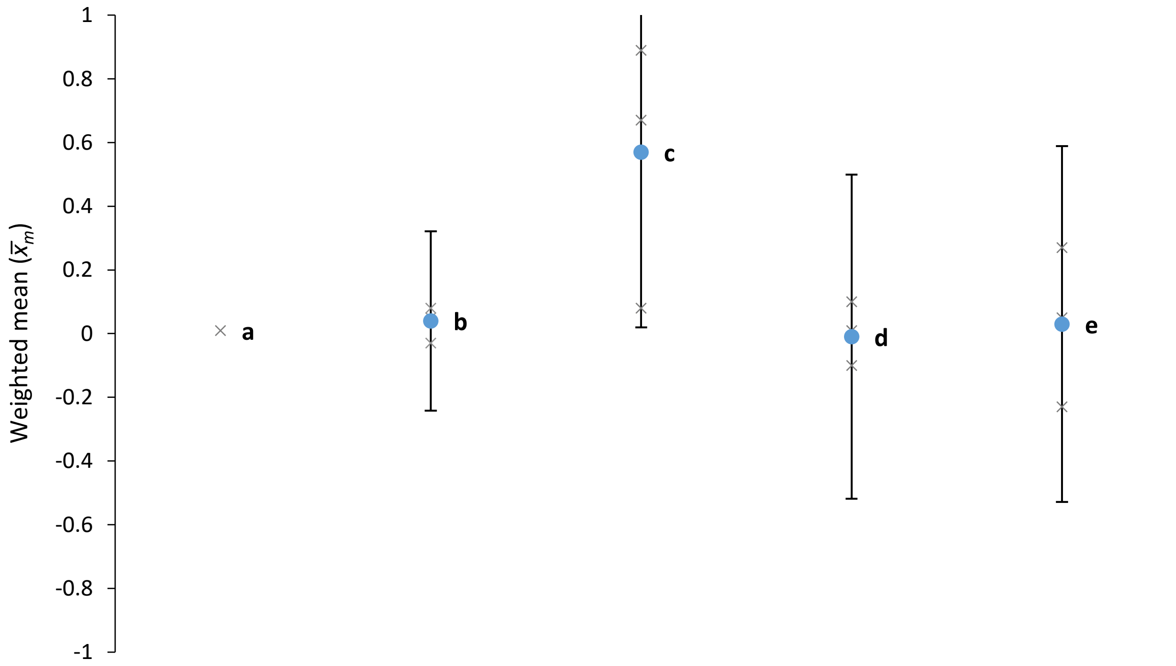
Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; e) Netherlands; f) South Korea; and g) USA.



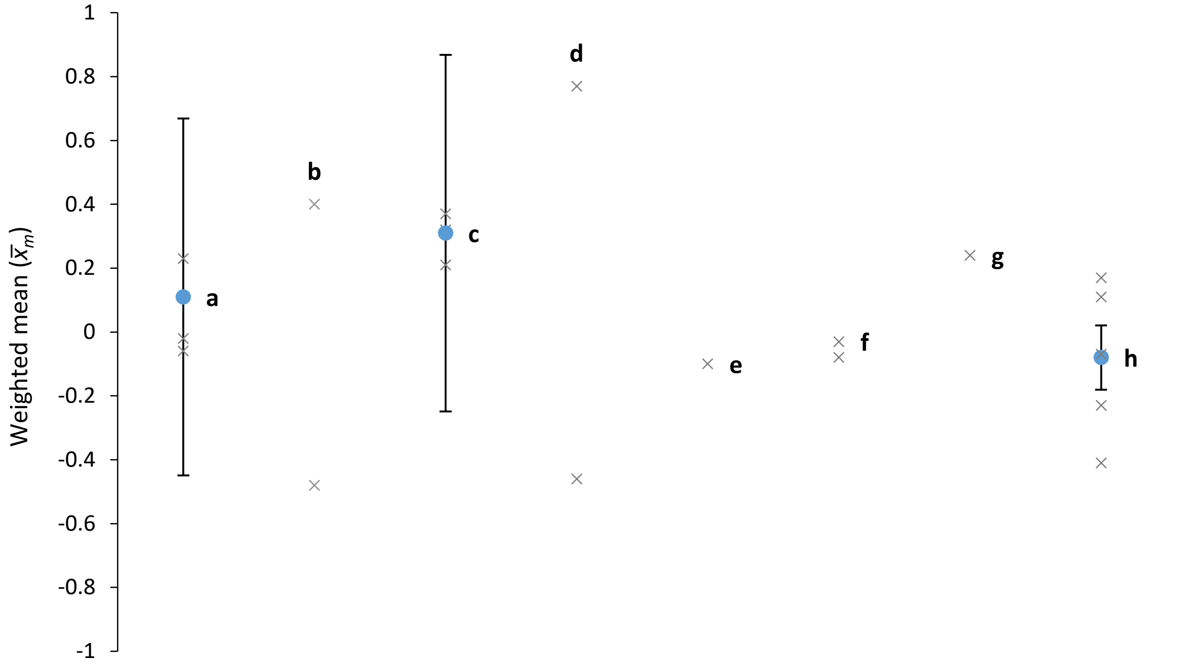
General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Germany; b) Italy; c) Netherlands; d) New Zealand; e) Taiwan; and f) USA.



Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Australia; b) France; c) Germany; d) Italy; e) Japan; f) Netherlands; g) South Korea; and h) USA.



Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; and e) Netherlands.



Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Australia; b) France; c) Germany; d) Italy; e) Japan; f) Netherlands; g) South Korea; and h) USA.

Supplementary Materials 2: Social Robot - Definition and Checklist

Definition of a Social Robot:

A social robot is a physically embodied artificial agent that: a) has design features which enable humans to perceive the agent as a social entity; b) is capable of interacting with humans via a social interface (Hegel, Muhl, Wrede, Hielscher-Fastabend, & Sagerer, 2009); c) can successfully communicate verbal and/or non-verbal information to humans.

In order for a social robot to be a physically embodied artificial agent, it needs to have a physical structure that mimics the behaviour, appearance, or movement of a living being (usually humans but also animals and plants). A robot can be considered to have a social interface if one of its purposes is engaging humans in social interaction. In short, a social robot is a system that can be perceived as a social entity that communicates with the user (Broekens, Heerink, & Rosendal, 2009).

Checklist used to operationalise the definition of a Social Robot:

1 Physical Embodiment

A social robot **must:**

1. Be physically embodied.

2. Have sensors capable of sensing, partially or fully, its operating environment as indicated by the presence of **at least two** of the following:
   - Camera, laser, sonar, or other vision system
   - Camera, laser, sonar, or other navigation system
   - Speech recognition system
   - Tactile sensors

3. Mimic, partially or fully, the behaviour of a living being (human, animal, or plant) by doing **at least one** of the following:
   - Mimic, partially or fully, the appearance of a living being (human, animal, or plant).
   - Mimic, partially or fully, the movement of a living being (human, animal, or plant).

2 Social Agency

A social robot **must be:**

1. Partially or fully, autonomous as indicated by **at least two** of the following:
    - Require little or no human input/ intervention to perform the task(s) it has been programmed to do.
    - Require little or no human input/ intervention to move, partially or fully, through its operating environment.
    - Require little or no human input/ intervention to sense its operating environment.

2. Able to identify other social agents (humans).

3. Identifiable by other social agents (humans) as a social entity as indicated by **at least three** of the following:
    - Ability of social agents to identify the robot by its physical structure.
    - Ability of social agents to approach the robot and engage it in interaction.
    - Ability of social agents to perceive the robot as an autonomous agent.
    - Ability of social agents to identify the robot's behaviour, appearance, or/and movement as, partially or fully, mimicking that of another living being (human, animal, or plant).

3 Social Interaction

A social robot **must:**

1. Have a social interface allowing the robot to engage and interact with humans in a social context as indicated by the presence of **at least two** of the following:
    - Speech recognition relevant to the robot's operational context.
    - Speech production relevant to the robot's operational context. *
    - Behaviour recognition relevant to the robot's operational context.
    - Behaviour production relevant to the robot's operational context. *

    AND **all** of the following:
    - Speech and/or behaviour production congruent, partially or fully, with human/animal social behaviour.
    - Speech and/or behaviour production and/or recognition that can be used to interact with other social agents (humans).

      ***** Other social agents must be able to, partially or fully, recognise and interpret robot speech and/or behaviour.

2. Be able to exchange verbal and/or non-verbal information with another social agent (human).

Supplementary Materials 3: Quality Assessment Tool

Studies were given a quality score between 1.0 (poor) and 4.0 (excellent). As the process of averaging is likely to produce decimals, quality scores were reported to one decimal point. Quality assessment relied on the accuracy of the information provided by authors (e.g., Cronbach's alpha) and the extent to which the review team could find evidence of quality (e.g., empirical studies supporting the validity of outcome measures). Any disagreements between review team members were resolved via discussion and consensus.

1 Study validity

1.1 Internal validity

Score	Are there any alternative plausible explanations (as far as the two review team members can detect) that could account for the results presented in the study?
1.0	There are one or more alternative plausible explanations due to one or more confounding variable(s); AND no attempt has been made to identify, explain, or otherwise account for these variables (e.g., no control or comparison group); AND alternative plausible explanations were neither considered nor discussed. OR The study has not been reported in sufficient detail to allow for a judgement to be made.
2.0	There are one or more alternative plausible explanations due to one or more confounding variable(s); AND some attempt has been made to identify, explain, or otherwise account for these variables; AND alternative plausible explanations were only discussed briefly and no modification to the conclusion was made to reflect this discussion.
3.0	There could be one or more alternative plausible explanations due to one or more confounding variable(s); AND some attempt has been made to identify, explain, or otherwise account for these variables; AND alternative plausible explanations were discussed briefly and the conclusion was modified to reflect this discussion.
4.0	There could be one or more alternative plausible explanations due to one or more confounding variable(s); AND an attempt has been made to comprehensively identify, explain, or otherwise account for these variables; AND alternative plausible explanations were discussed in detail and the conclusion was modified to reflect this discussion.

OR It is unlikely that there are any alternative plausible explanations.

1.2 External validity

Score	(a) Is there any evidence* of sampling bias**? * based on the information provided by the study's author(s) ** sampling bias to mean any factor or procedure (intended or unintended) that leads to the selection of an unrepresentative of the target population sample. Leniency (plus 0.5 to score) was shown to studies which clearly identify any sampling bias and attempt to adjust their conclusions/analysis as a result.
1.0	A non-probability sampling method was used to select participants (e.g., volunteers); OR Sampling procedure has not been explained; OR A probability sampling method has been used but is inappropriate to answer the research question (e.g., stratified sampling used when partitioning of the population into groups is not appropriate for the research question).
2.0	A probability sampling method has been used but there is doubt as to whether the sampling procedure has been carried out correctly (e.g., reported stratified sampling but unclear whether a simple random sample has been obtained from each group).
3.0	A probability sampling method (stratified sampling, cluster sampling, systematic sampling, or combination) has been used and there is little or no doubt as to whether the sampling procedure has been carried out correctly.
4.0	Simple random sampling has been used and there is little or no doubt as to whether the sampling procedure has been carried out correctly.

Score	**(b) How representative is the sample of the target population*?** * target population as defined by the authors or indicated in the research question(s) or hypotheses
1.0	Entire sample appears to be completely unrepresentative of the target population; OR Entire sample appears to represent a minority or atypical subgroup of the target population; OR Sample has not been described in sufficient detail to make a judgement.
2.0	A large portion of the sample appears to be completely unrepresentative of the target population; OR A large portion of the sample appears to represent a minority or atypical subgroup of the target population.
3.0	A small portion of the sample appears to be completely unrepresentative of the target population; OR A small portion of the sample appears to represent a minority or atypical subgroup of the target population.
4.0	Entire sample appears to be mostly or completely representative of the target population.

2 Outcome measures

2.1 Validity

Score	**(a) Has this measure been used* in other studies investigating attitudes toward, trust in, acceptance of, or anxiety toward robots?** * judgement of the prior use of the measure was first guided by the information provided by the authors of the study under assessment (e.g., if authors provide evidence of multiple use of the measure by different authors, a score of 4 will be given). If authors provided no explicit information regarding prior use, an effort was made to check whether the measure has been used before. No penalty was applied to studies failing to evidence prior use.
1.0	Measure was developed specifically for the study and has not been used previously.
2.0	Measure has been used previously in multiple (two or more) studies to measure something other than attitudes, trust, acceptance, or anxiety toward robots; OR Measure has been used previously in only one other study by the same authors.
3.0	Measure has been used previously in multiple (two or more) studies by the same authors to measure attitudes, trust, acceptance, or anxiety toward robots.
4.0	Measure has been used previously in multiple (two or more) studies by different authors to measure attitudes, trust, acceptance, or anxiety toward robots.

Score	(b) What evidence* is there for the validity of the measure? Does it measure attitudes toward, trust in, acceptance of, or anxiety toward robots**? * here evidence means an empirical study with the explicit aim to test at least one aspect of the measure's validity. Multiple empirical studies may be published in the same paper but were counted individually. ** measures developed and validated specifically in the context of measuring attitudes toward, trust in, acceptance of, and anxiety toward robots was given a higher score than similarly validated measures in a different context (e.g., a measure of anxiety toward robots will be rated higher than a similar measure of anxiety toward humans).
1.0	No attempts have been made to assess the validity of the measure; OR Any attempts to assess the validity of the measure are inadequate or inappropriate; OR Existing empirical evidence does not support the validity of the measure.
2.0	Some evidence (at least one empirical study) is available but only supports some types of the measure's validity; OR Any evidence supporting the validity of the measure is not in the context of attitudes toward, trust in, acceptance of, or anxiety toward robots.
3.0	Some evidence (at least one empirical study) is available to support the validity of the measure; OR substantial evidence (three or more empirical studies) is available but only supports some types of the measure's validity.
4.0	Substantial evidence (three or more empirical studies) is available to support the validity of the measure.

2.2 Reliability

Score	(a) What evidence is there for the test-retest reliability of the measure?
1.0	Test-retest reliability was measured and reported adequately within the study, the measure's reliability was poor (correlation of 0.5 > *r*); OR An attempt has been made to assess the test-retest reliability of the measure within the study but reliability was measured and/or reported inadequately; OR No attempts have ever been made to assess the test-retest reliability of the measure.
2.0	Test-retest reliability was measured and reported adequately within the study, the measure's reliability was questionable (correlation of 0.7 > $r \geq$ 0.5); OR Test-retest reliability was previously measured and reported adequately within a different study, the measure's reliability was questionable-good (correlation of 0.8 > $r \geq$ 0.6).
3.0	Test-retest reliability was measured and reported adequately within the study, the measure's reliability was good (correlation of 0.8 > $r \geq$ 0.7); OR test-retest reliability was previously measured and reported adequately within a different study, the measure's reliability was good-excellent (correlation of 1 > $r \geq$ 0.7).
4.0	Test-retest reliability was measured and reported adequately within the study, the measure's reliability was excellent (correlation of $r \geq$ 0.8).

Score	(b) What evidence is there for the internal consistency reliability of the measure (as defined by Cronbach's alpha)?
1.0	Internal consistency reliability was measured and reported adequately within the study, the measure's reliability was poor (coefficient of $0.5 > \alpha$); OR An attempt has been made to assess the internal consistency reliability of the measure within and/or outside of the study but reliability was measured and/or reported inadequately; OR No attempts have ever been made to assess the internal consistency reliability of the measure.
2.0	Internal consistency reliability was measured and reported adequately within the study, the measure's reliability was questionable (coefficient of $0.7 > \alpha \geq 0.5$); OR Internal consistency reliability was previously measured and reported adequately within a different study, the measure's reliability was questionable-good (coefficient of $0.8 > \alpha \geq 0.6$).
3.0	Internal consistency reliability was measured and reported adequately within the study, the measure's reliability was good (coefficient of $0.8 > \alpha \geq 0.7$); OR internal consistency reliability was previously measured and reported adequately within a different study, the measure's reliability was good-excellent (coefficient of $1 > \alpha \geq 0.7$).
4.0	Internal consistency reliability was measured and reported adequately within the study, the measure's reliability was excellent (coefficient of $\alpha \geq 0.8$).

3 Objectivity

Score	**How objective* is the measure of attitudes toward, trust in, acceptance of, or anxiety toward robots?** * objective to mean something that is externally observable and verifiable and its measurement is not dependent on mental or subjective personal experience (although it may be affected by it).
1.0	Data collected using this measure is assumed to represent participants' self-reported internal states (e.g., beliefs); AND any analysis and subsequent conclusions derived from the collected data are subject to the interpretation of the researcher (e.g., discourse analysis of qualitative data from interviews or focus groups).
2.0	Data collected using this measure is assumed to represent participants' self-reported internal states (e.g., beliefs) AND data can be quantified to allow for statistical analysis and subsequent interpretation by the researcher (e.g., ANOVA analysis of Likert scale items in a questionnaire).
3.0	Data collected using this measure is not self-reported but still assumed to represent participants' internal states (e.g., attitudes) to some extent; AND data is considered at less risk of response bias (e.g., social desirability) but could still be influenced knowingly by the participant in some cases; AND data is inherently quantitative (e.g., reaction time, duration of eye gaze) and allows for statistical analysis and subsequent interpretation by the researcher.
4.0	Data collected using this measure is not self-reported but still assumed to represent participants' internal states (e.g., anxiety) to some extent; AND data is considered at almost no risk of response bias (e.g., social desirability) and is unlikely to be knowingly influenced by the participant; AND data is inherently quantitative (e.g., heart rate, skin conductance, pupil dilatation) and allows for statistical analysis and subsequent interpretation by the researcher.

Supplementary Materials 4: Keywords that the robots were programmed to detect

Studies 2 and 3 (shampoo)	Study 4 (films)
recommend	recommendation
recommendations	recommendations
advise	advise
explain	explain
properties	specific film
specific product	films
product	what have you got
shampoo	Can't hardly wait
shampoos	Judgment Night
the oil control shampoo	Misiss Brown
the repairing shampoo	The Browning Version
the color protect shampoo	Asylum
the thickening shampoo	The secret of Kells
the shine shampoo	Hunt for the Wilderpeople
the anti-dandruf shampoo	Hardware
the volumizing shampoo	comedy
the curls and waves shampoo	action
oily	romance
oily hair	drama
dry	horror
dry hair	animation
damaged	adventure
damaged hair	sci-fi
broken	science fiction
split ends	I don't know
dyed	any

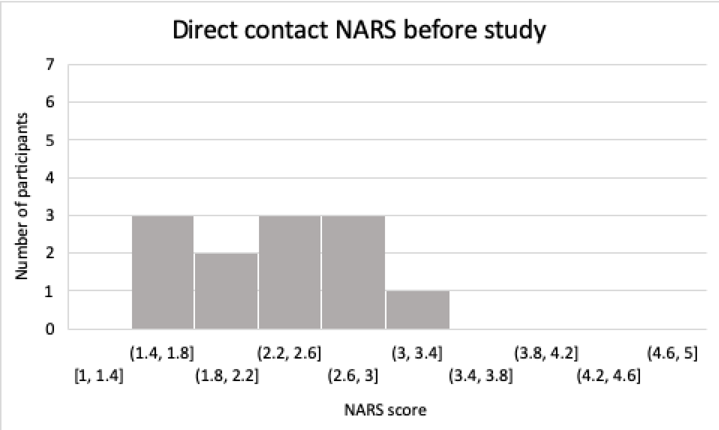
dyed hair	yes
colored hair	yes please
colored	please
colorful	ok
colors	right
thin	all right
thin hair	yep
weak	yeah
straight	ya
normal	no
normal hair	na
grey	no thanks
grey hair	
shine	
long	
long hair	
short	
short hair	
dandruff	
curly	
curly hair	
wavy	
wavy hair	
I don’t know	
yes	
yes please	
please	
ok	
right	

all right yep yeah ya no na no thanks	

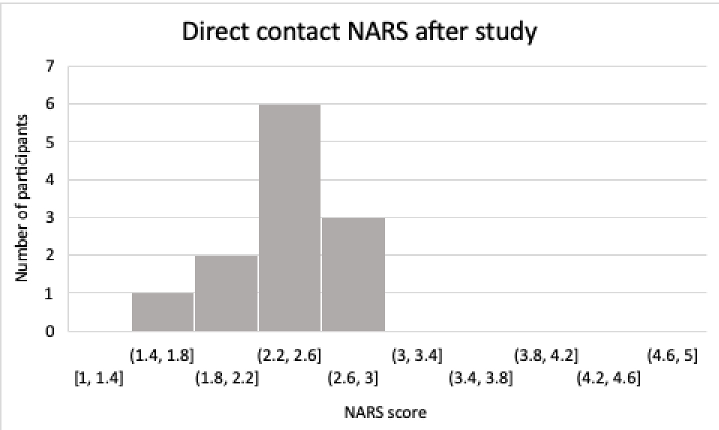
Supplementary Materials 5: Histograms representing the data in study 2

The following histograms show descriptively the data collected in study 2. These data are separated according to condition and dependent variable.

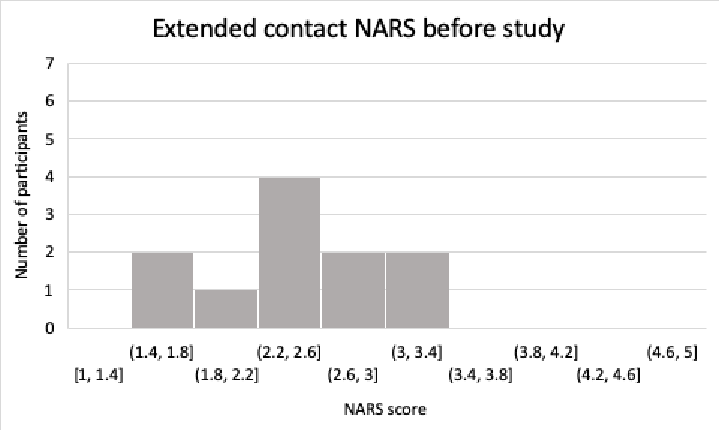
<u>Explicit attitudes: NARS</u>



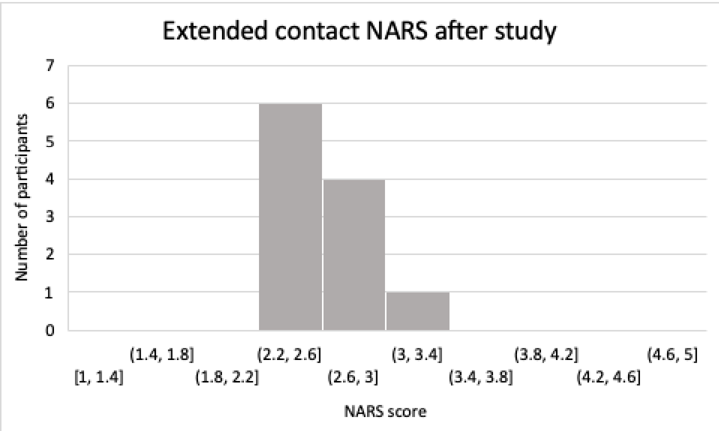
Participants' explicit attitudes towards robots in the direct contact condition in Study 2 before they had direct contact with the robot Pepper (*M=2.32*, *SD=.55*).



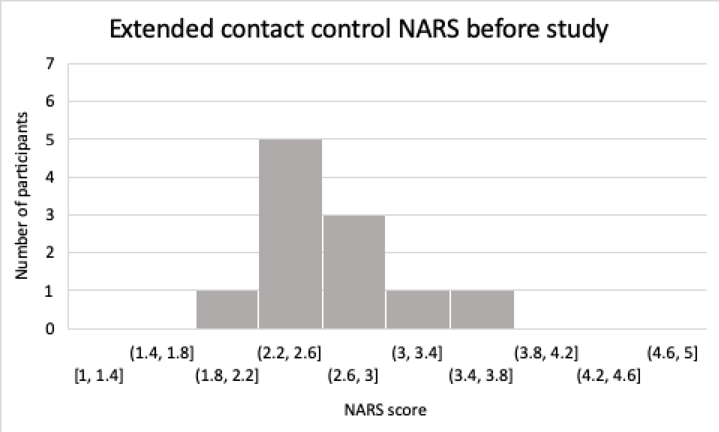
Participants' explicit attitudes towards robots in the direct contact condition in Study 2 after they had direct contact with the robot Pepper (*M=2.38*, *SD=.37*).



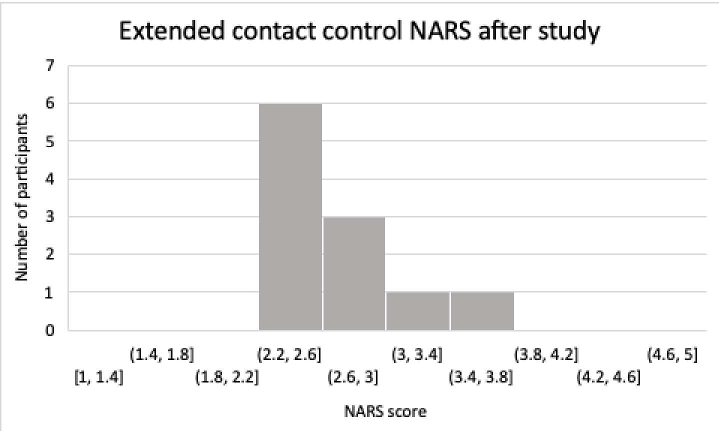
Participants' explicit attitudes towards robots in the extended contact condition in Study 2 before they watched the video (*M=2.49, SD=.51*).



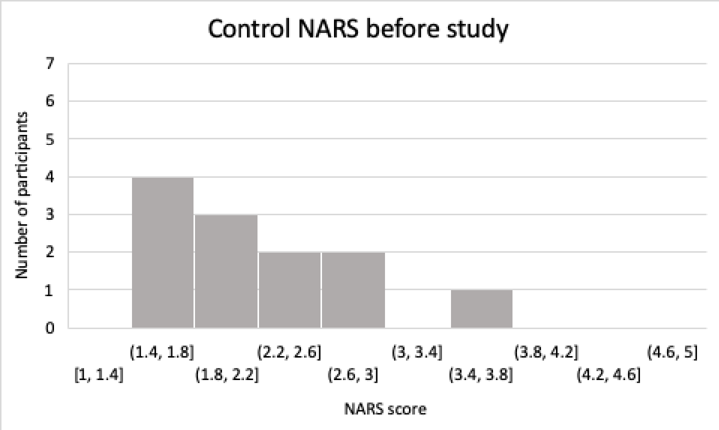
Participants' explicit attitudes towards robots in the extended contact condition in Study 2 after they watched the video (*M=2.58*, *SD=.24*).



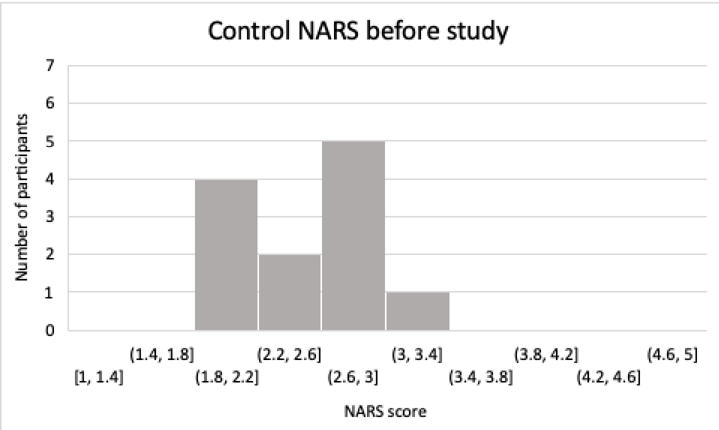
Participants' explicit attitudes towards robots in the extended contact control condition in Study 2 before they watched the video (*M=2.70, SD=.46*).



Participants' explicit attitudes towards robots in the extended contact condition in Study 2 after they watched the video

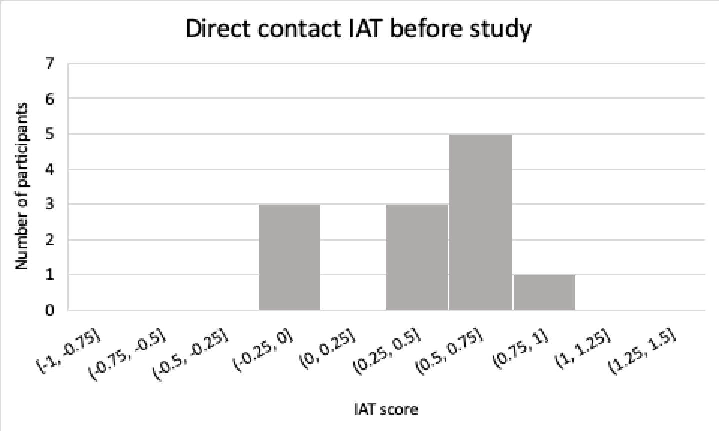


Participants' explicit attitudes towards robots in the control condition in Study 2 before they performed the experiment (*M=2.23*, *SD=.65*).

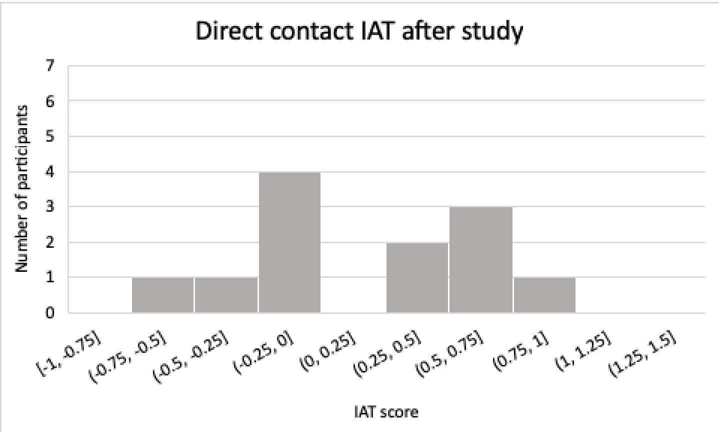


Participants' explicit attitudes towards robots in the control condition in Study 2 after they performed the experiment (*M=2.53*, *SD=.45*).

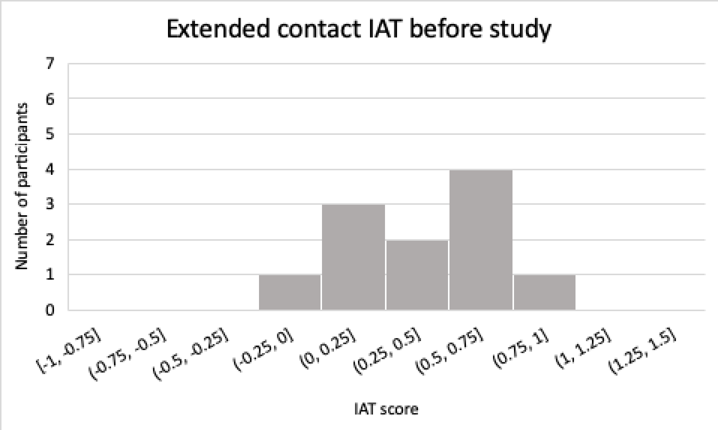
Implicit attitudes: IAT



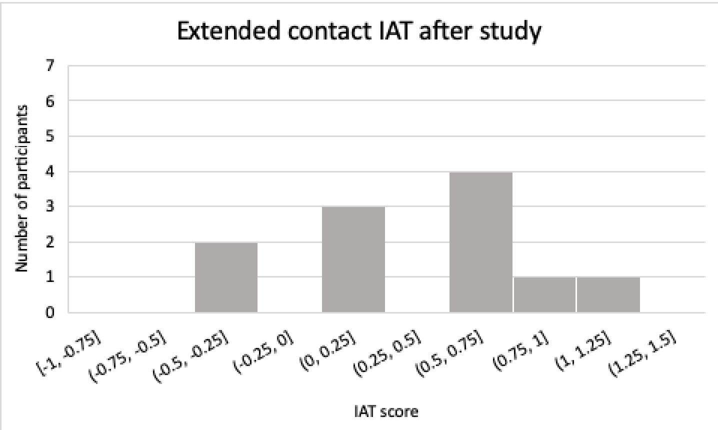
Participants' implicit attitudes towards robots in the direct contact condition in Study 2 before they had direct contact with the robot Pepper (*M=.41*, *SD=.33*).



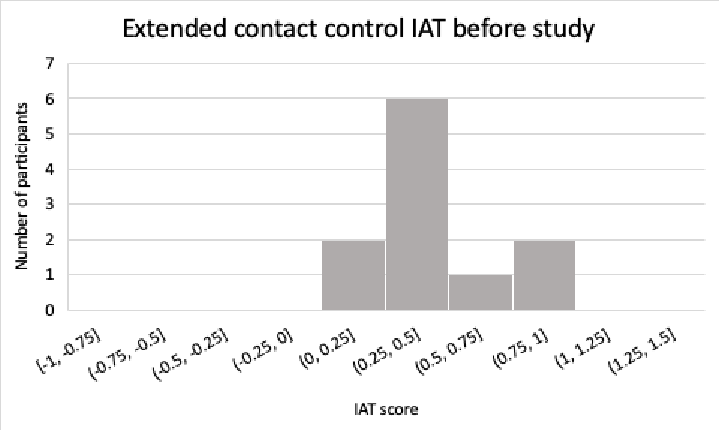
Participants' implicit attitudes towards robots in the direct contact condition in Study 2 after they had direct contact with the robot Pepper (*M=.20*, *SD=.45*).



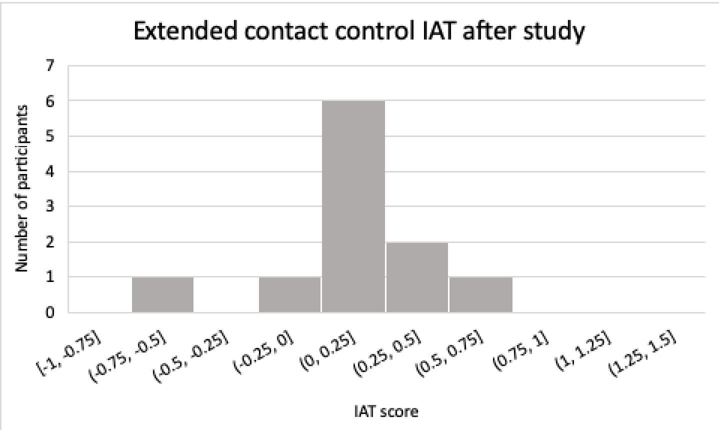
Participants' implicit attitudes towards robots in the extended contact condition in Study 2 before they watched the video (*M=.44, SD=.33*).



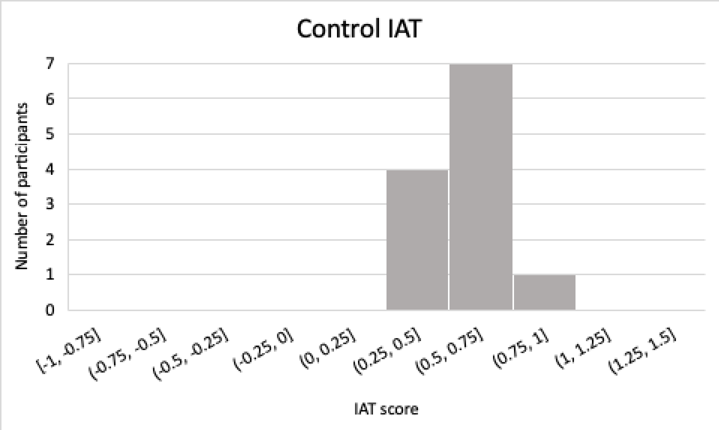
Participants' implicit attitudes towards robots in the extended contact condition in Study 2 after they watched the video (*M=.35, SD=.50*).



Participants' implicit attitudes towards robots in the extended contact control condition in Study 2 before they watched the video (*M=.48, SD=.24*).

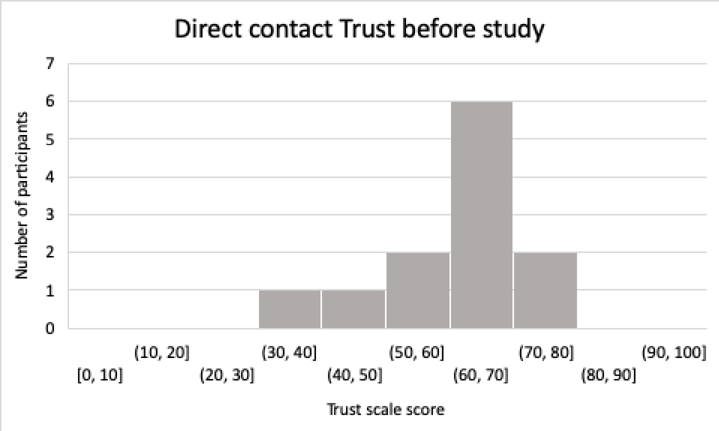


Participants' implicit attitudes towards robots in the extended contact control condition in Study 2 after they watched the video (*M=.14, SD=.33*).

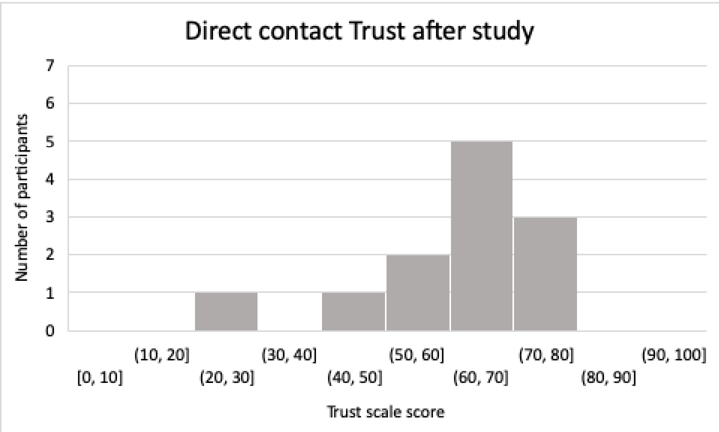


Participants' implicit attitudes towards robots in the control condition in Study 2 (*M=.56, SD=.13*).

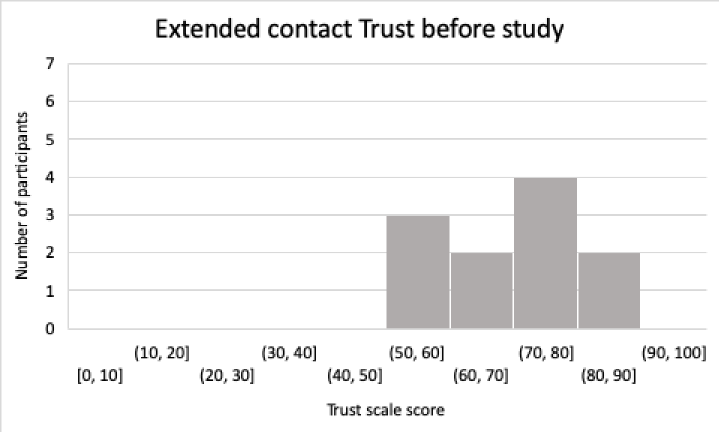
Trust: 40 items Trust scale



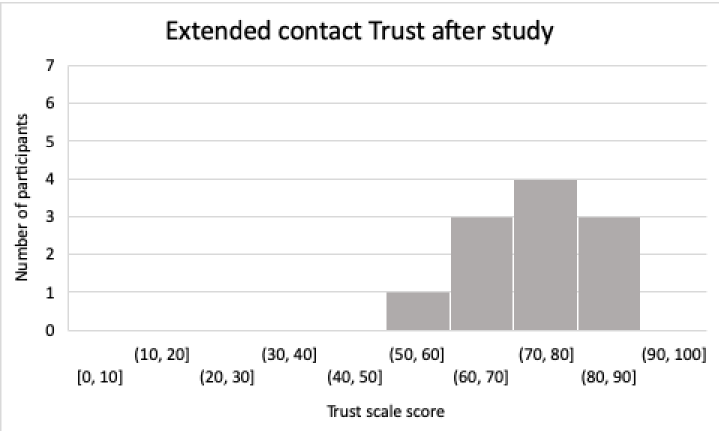
Participants' trust in robots in the direct contact condition in Study 2 before they had direct contact with the robot Pepper (*M=61.24, SD=10.42*).



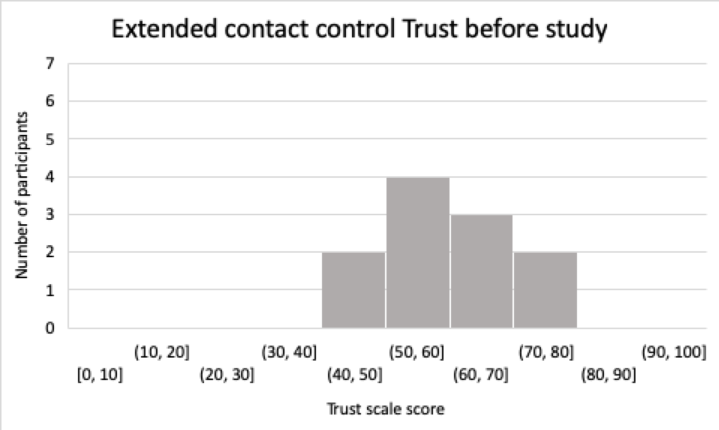
Participants' trust in robots in the direct contact condition in Study 2 after they had direct contact with the robot Pepper (*M=62.38, SD=14.15*).



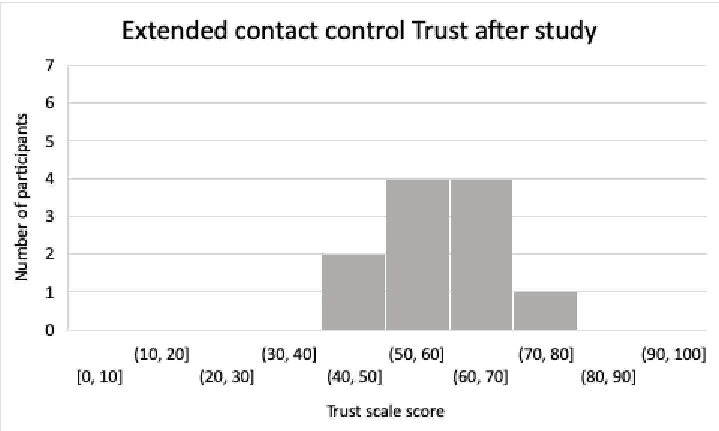
Participants' trust in robots in the extended contact condition in Study 2 before they watched the video (*M=70.04, SD=10.29*).



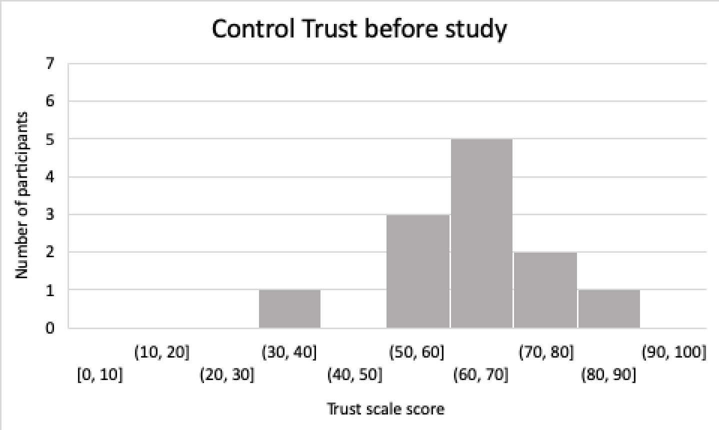
Participants' trust in robots in the extended contact condition in Study 2 after they watched the video (*M=73.53, SD=10.41*).



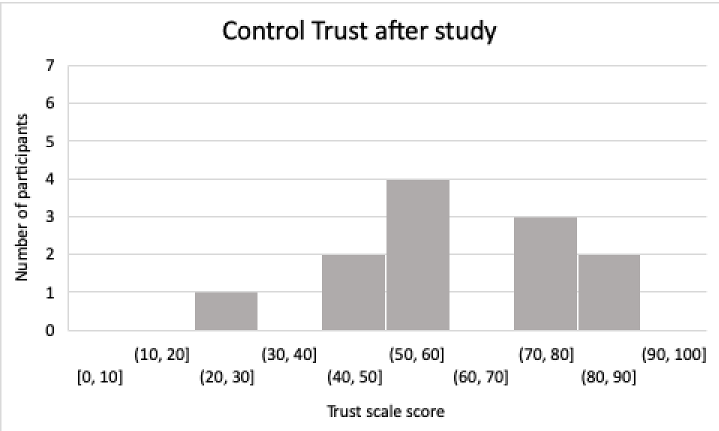
Participants' trust in robots in the extended contact control condition in Study 2 before they watched the video (*M=59.53*, *SD=9.53*).



Participants' trust in robots in the extended contact control condition in Study 2 after they watched the video (*M=59.04*, *SD=7.87*).



Participants' trust in robots in the control condition in Study 2 before they performed the experiment (*M=62.79, SD=13.95*).

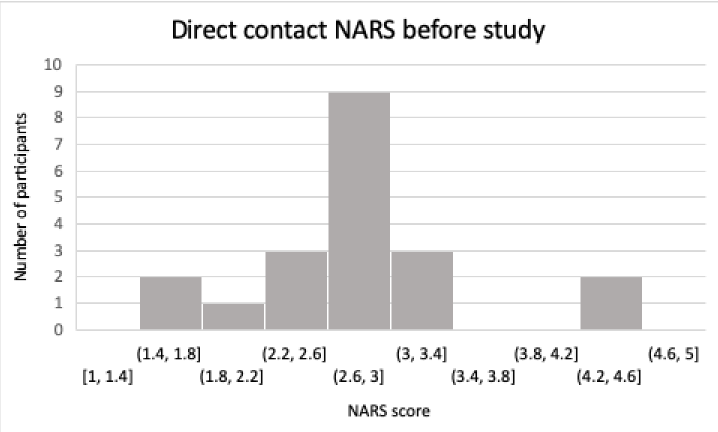


Participants' trust in robots in the control condition in Study 2 after they performed the experiment (*M=62.79, SD=13.95*).

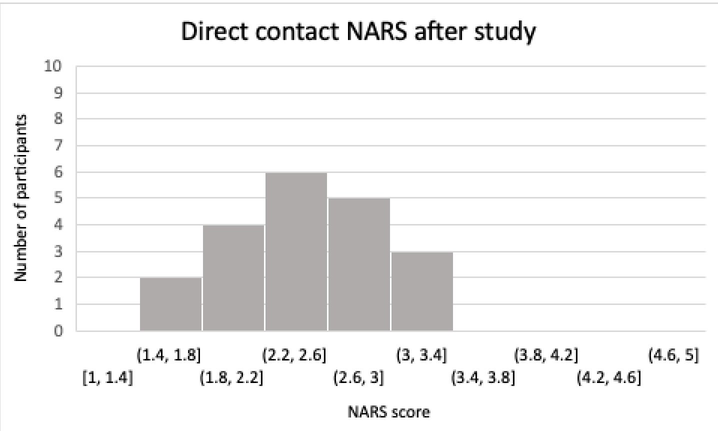
Supplementary Materials 6: Histograms representing the data in study 3

The following histograms show descriptively the data collected in study 3. These data are separated according to condition and dependent variable.

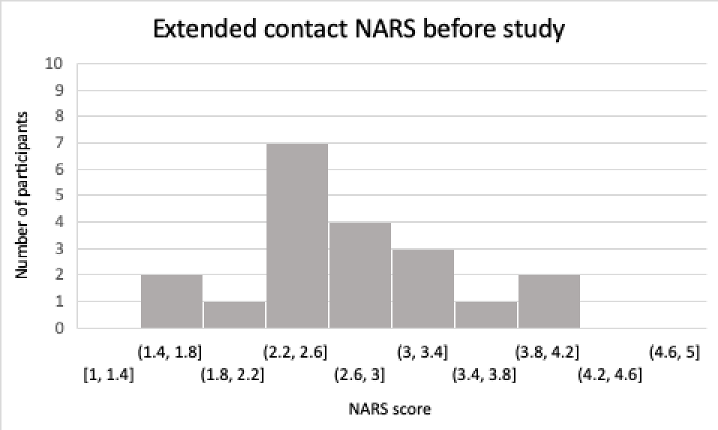
<u>Explicit attitudes: NARS</u>



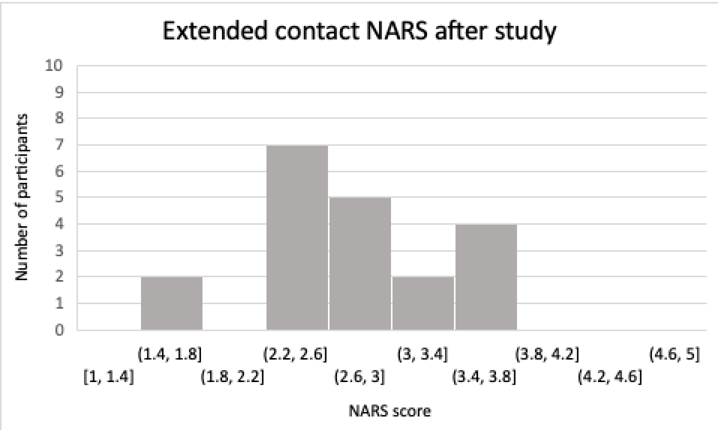
Participants' explicit attitudes towards robots in the direct contact condition in Study 3 before they had direct contact with the robot Pepper (*M=2.81*, *SD=.70*).



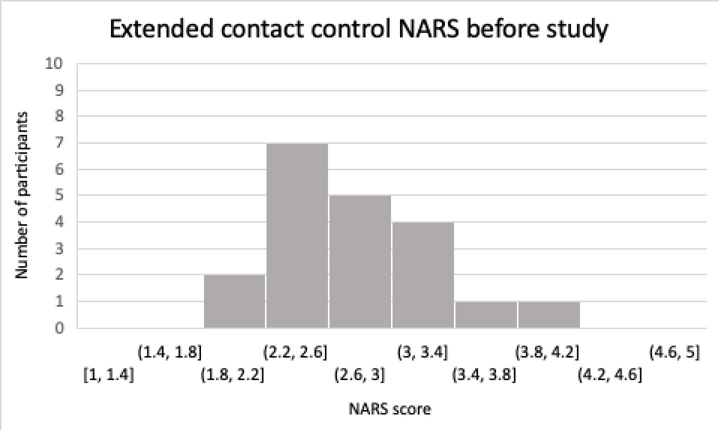
Participants' explicit attitudes towards robots in the direct contact condition in Study 3 after they had direct contact with the robot Pepper (*M=2.48, SD=.50*).



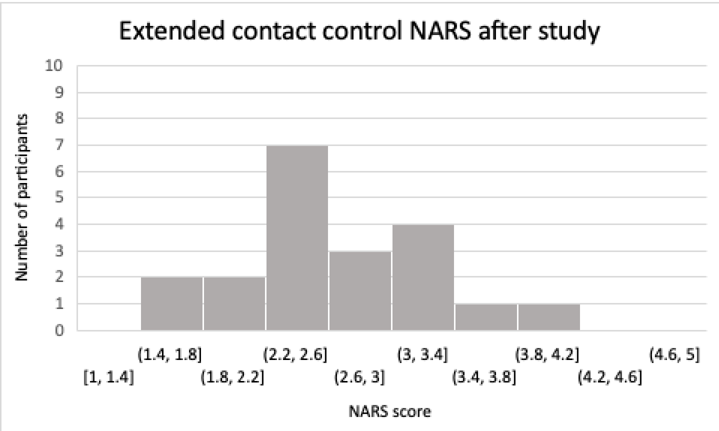
Participants' explicit attitudes towards robots in the extended contact condition in Study 3 before they watched the video (*M=2.75, SD=.65*).



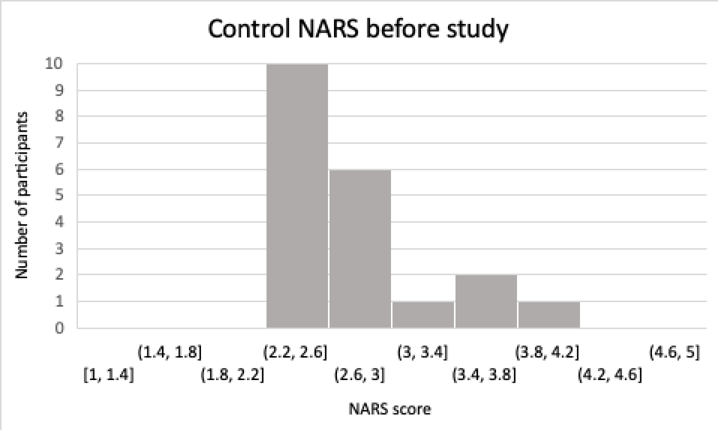
Participants' explicit attitudes towards robots in the extended contact condition in Study 3 after they watched the video (*M=2.74, SD=.63*).



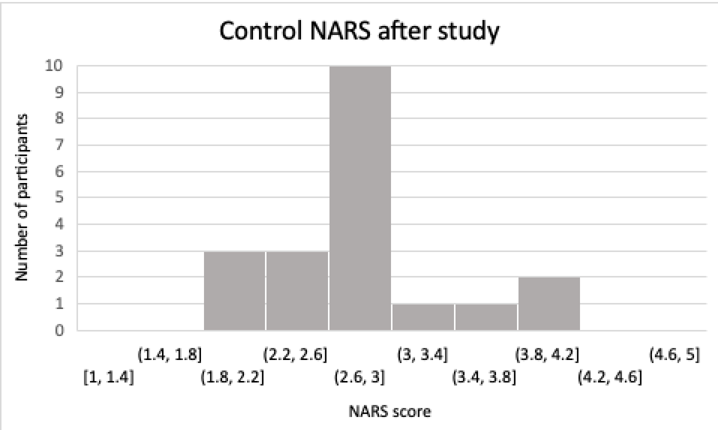
Participants' explicit attitudes towards robots in the extended contact control condition in Study 3 before they watched the video (*M=2.72, SD=.53*).



Participants' explicit attitudes towards robots in the extended contact control condition in Study 3 after they watched the video (*M=2.64, SD=.62*).

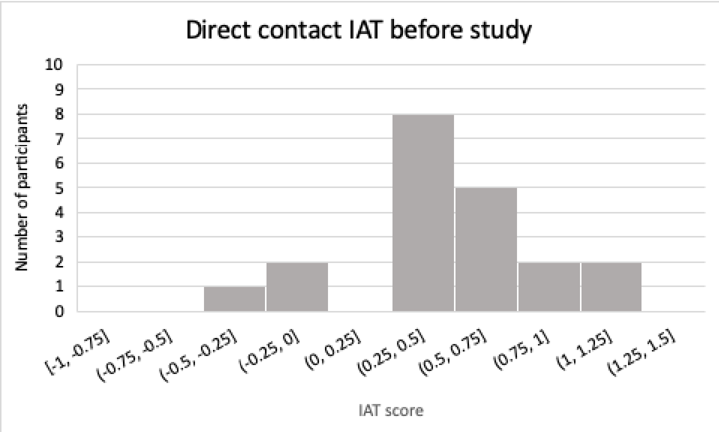


Participants' explicit attitudes towards robots in the control condition in Study 3 before they performed the experiment (*M=2.80, SD=.60*).

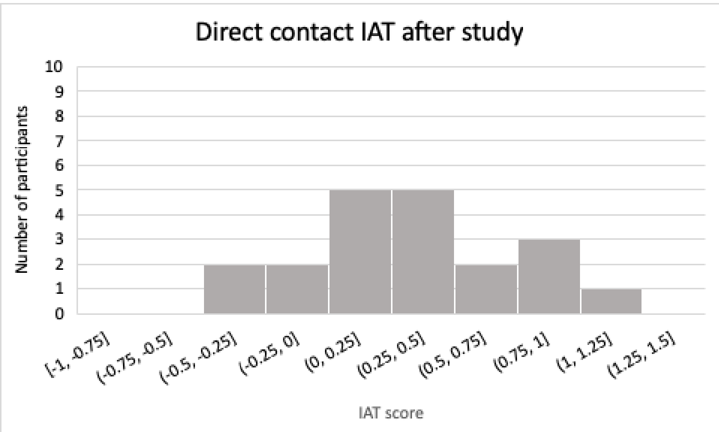


Participants' explicit attitudes towards robots in the control condition in Study 3 after they performed the experiment (*M=2.79, SD=.57*).

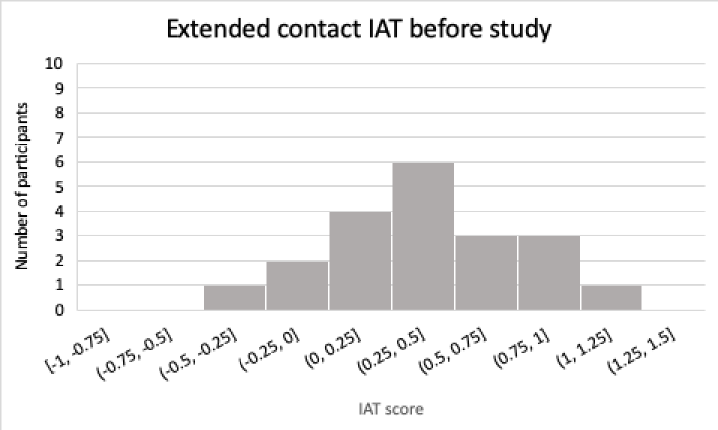
Implicit attitudes: IAT



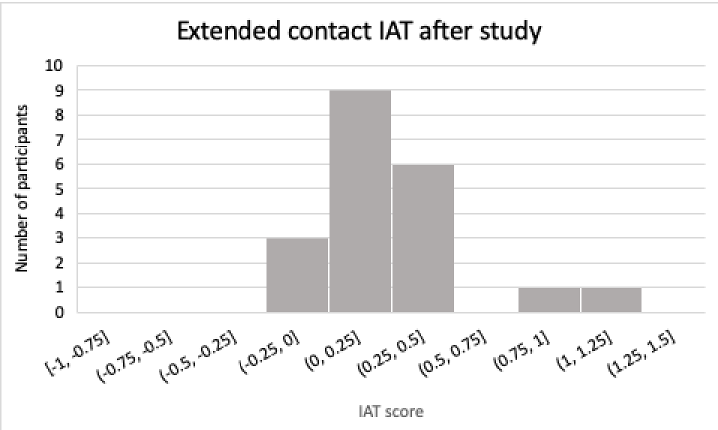
Participants' implicit attitudes towards robots in the direct contact condition in Study 3 before they had direct contact with the robot Pepper (*M=.50*, *SD=.35*).



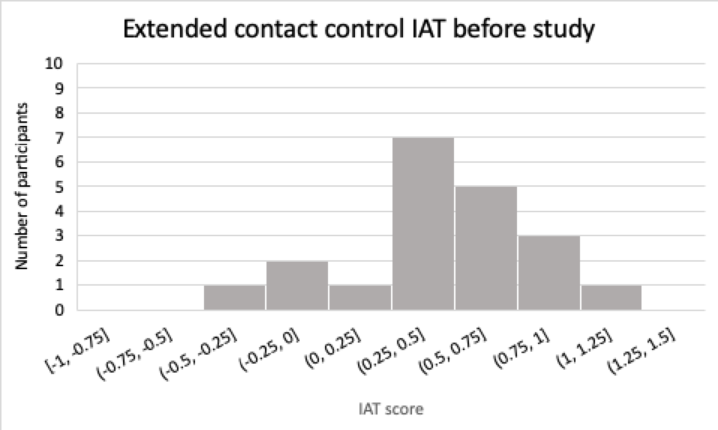
Participants' implicit attitudes towards robots in the direct contact condition in Study 3 after they had direct contact with the robot Pepper (*M=.31*, *SD=.38*).



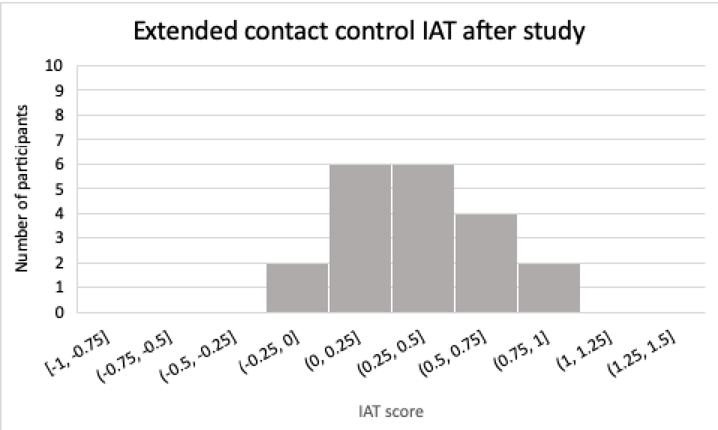
Participants' implicit attitudes towards robots in the extended contact condition in Study 3 before they watched the video (*M=.40, SD=.36*).



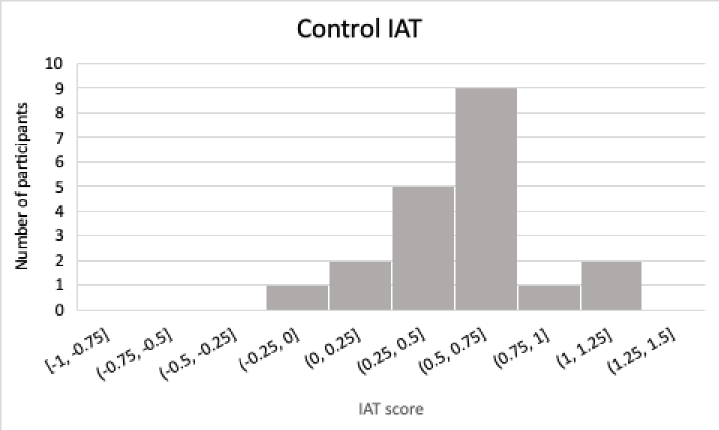
Participants' implicit attitudes towards robots in the extended contact condition in Study 3 after they watched the video (*M=.25, SD=.32*).



Participants' implicit attitudes towards robots in the extended contact control condition in Study 3 before they watched the video (*M=.44, SD=.36*).

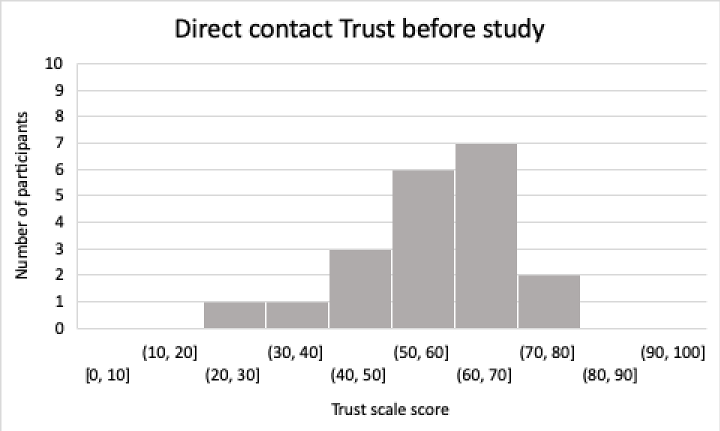


Participants' implicit attitudes towards robots in the extended contact control condition in Study 3 after they watched the video (*M=.34*, *SD=.30*).

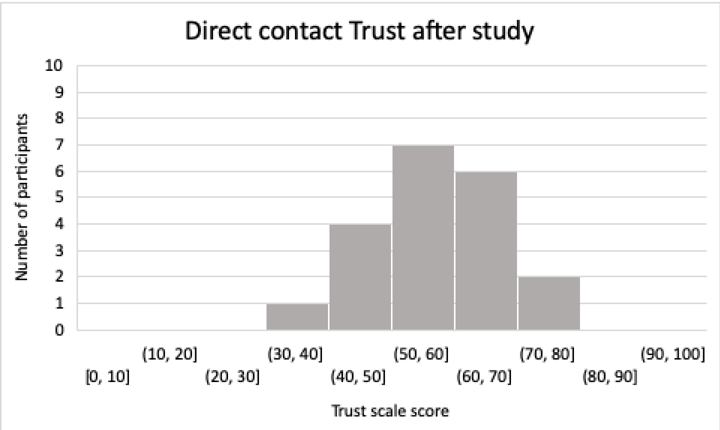


Participants' implicit attitudes towards robots in the control condition in Study 3 (*M=.54*, *SD=.30*).

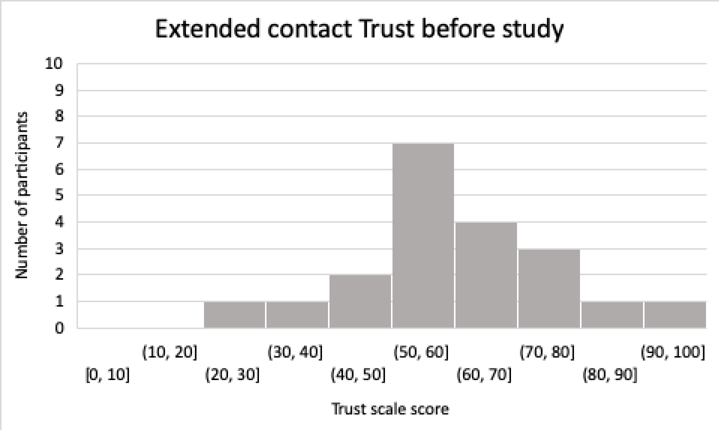
Trust: 40 items Trust scale



Participants' trust in robots in the direct contact condition in Study 3 before they had direct contact with the robot Pepper (*M=56.55*, *SD=11.42*).

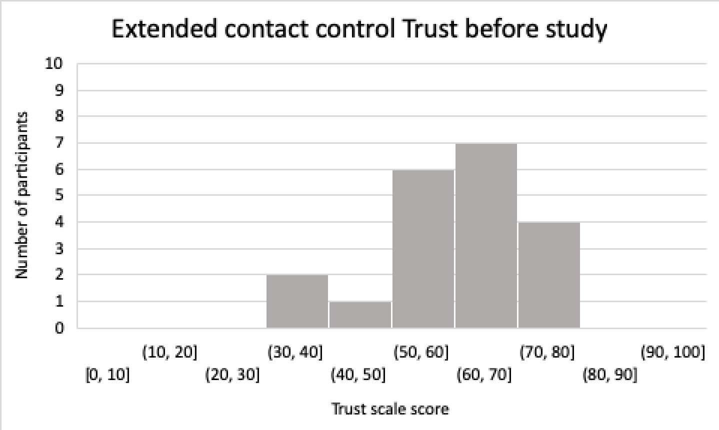


Participants' trust in robots in the direct contact condition in Study 3 after they had direct contact with the robot Pepper (*M=58.41*, *SD=10.00*).

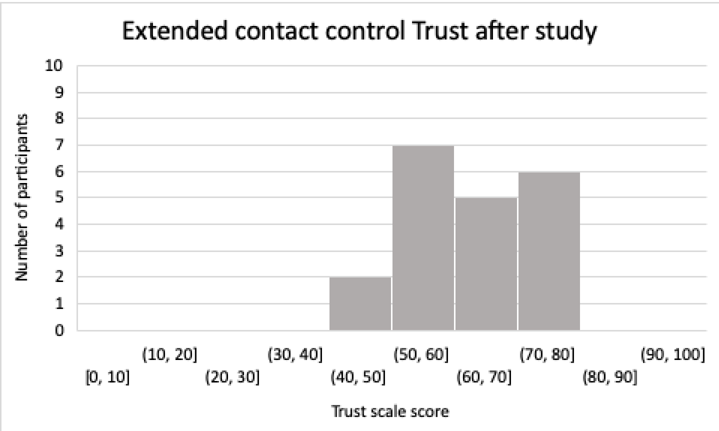


Participants' trust in robots in the extended contact condition in Study 3 before they watched the video (*M=59.17*, *SD=15.00*).

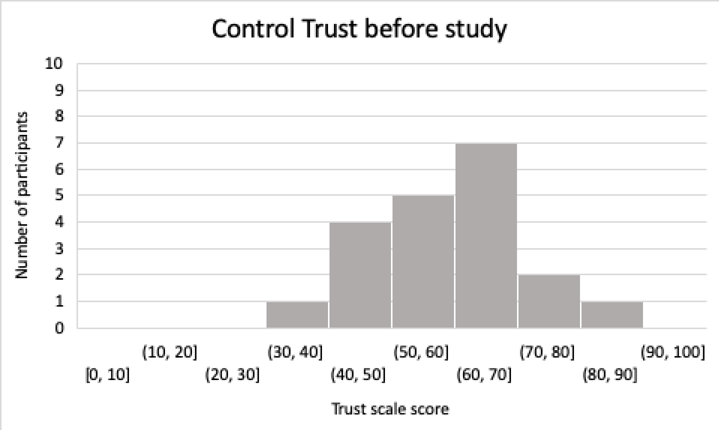
Participants' trust in robots in the extended contact condition in Study 3 after they watched the video (*M=61.79, SD=11.74*).



Participants' trust in robots in the extended contact control condition in Study 3 before they watched the video (*M=59.90, SD=11.71*).



Participants' trust in robots in the extended contact control condition in Study 3 after they watched the video (*M=61.70, SD=11.41*).



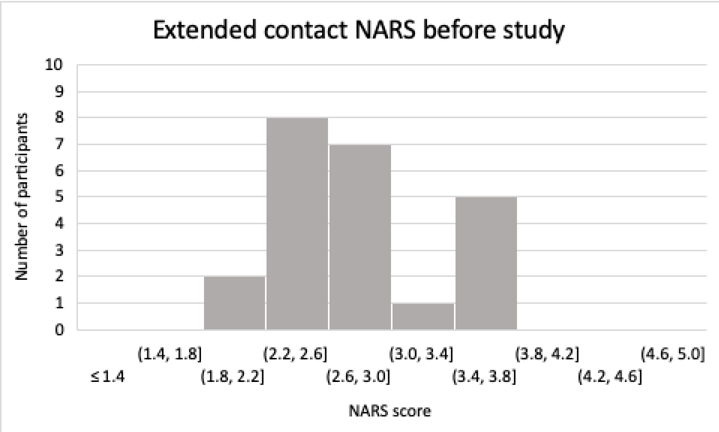
Participants’ trust in robots in the control condition in Study 3 before they performed the experiment (*M=59.08, SD=11.60*).

Participants’ trust in robots in the control condition in Study 3 after they performed the experiment (*M=62.18, SD=12.78*).

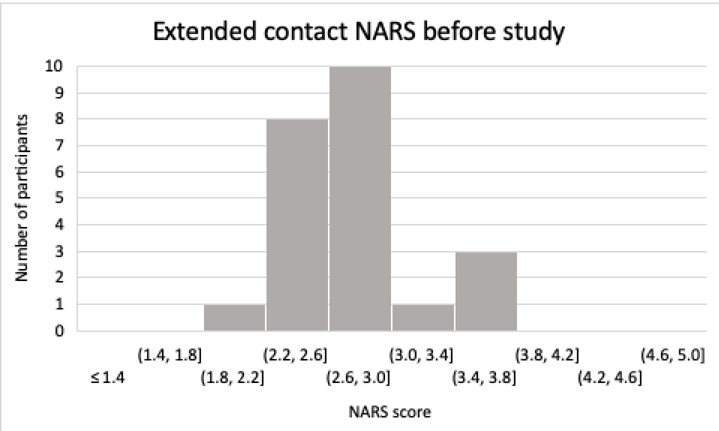
Supplementary Materials 7: Histograms representing the data in study 4

The following histograms show descriptively the data collected in study 4. These data are separated according to condition and dependent variable.

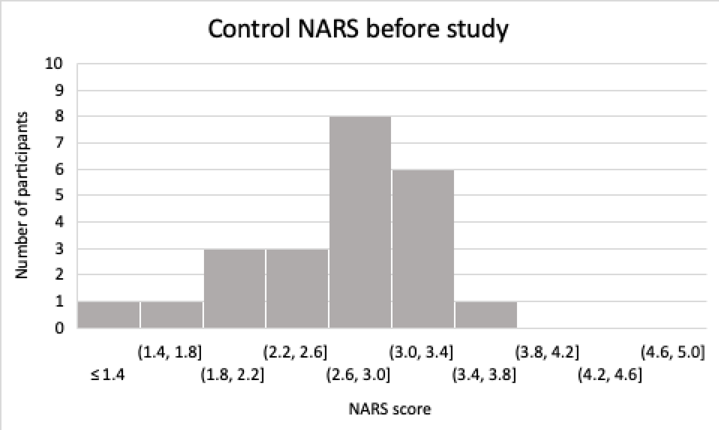
<u>Explicit attitudes: NARS</u>



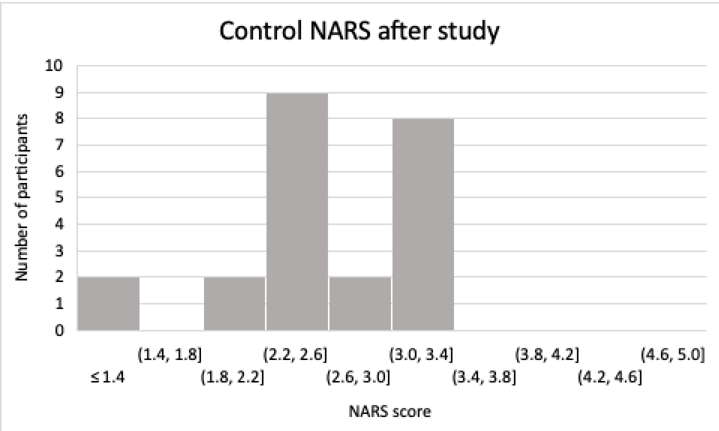
Participants' explicit attitudes towards robots in the extended contact condition in Study 4 before they watched the video (*M=2.79*, *SD=.48*).



Participants' explicit attitudes towards robots in the extended contact condition in Study 4 after they watched the video (*M=2.76*, *SD=.42*).

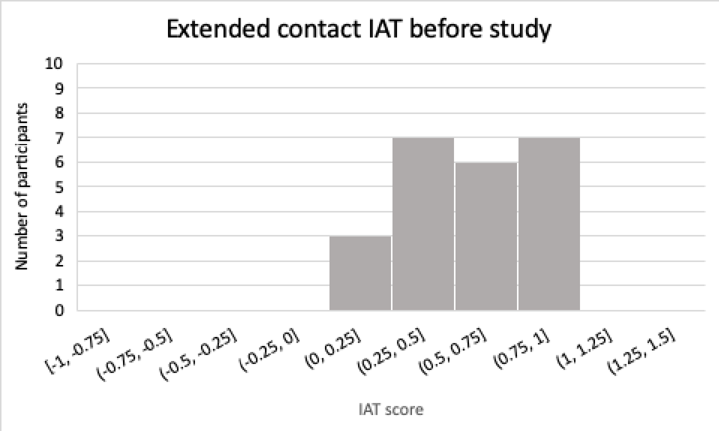


Participants' explicit attitudes towards robots in the control condition in Study 4 before they performed the experiment (*M=2.65, SD=.59*).

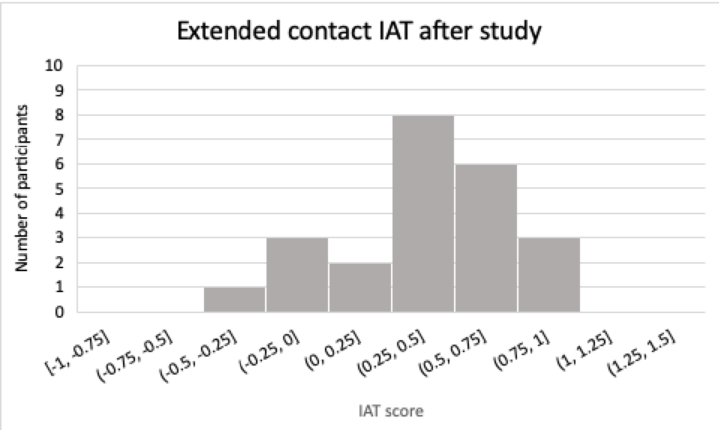


Participants' explicit attitudes towards robots in the control condition in Study 4 after they performed the experiment (*M=2.57, SD=.60*).

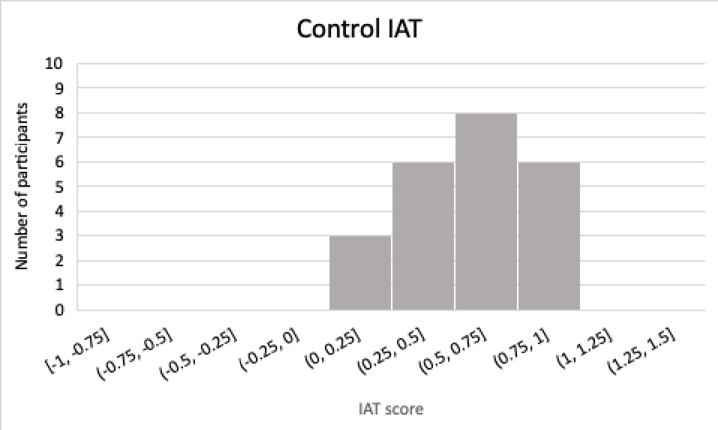
Implicit attitudes: IAT



Participants' implicit attitudes towards robots in the extended contact condition in Study 4 before they watched the video (*M=.56, SD=.25*).

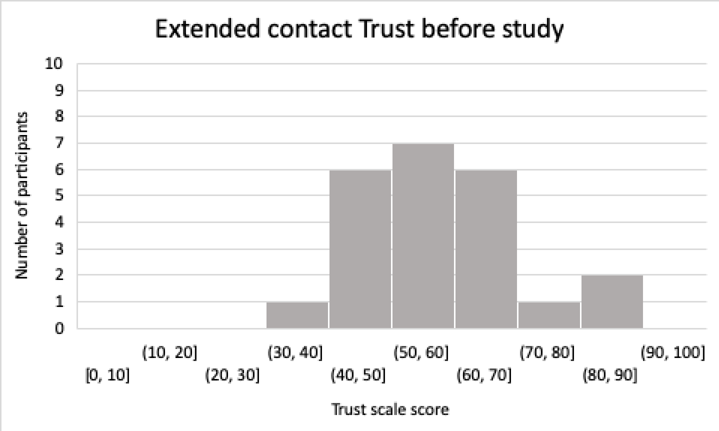


Participants' implicit attitudes towards robots in the extended contact condition in Study 4 after they watched the video (*M=.38, SD=.34*).



Participants' implicit attitudes towards robots in the control condition in Study 4 (*M=.55, SD=.23*).

Trust: 40 items Trust scale



Participants’ trust in robots in the extended contact condition in Study 4 before they watched the video (*M=57.09, SD=12.76*).

Participants’ trust in robots in the extended contact condition in Study 4 after they watched the video (*M=61.00, SD=12.81*).

Participants’ trust in robots in the control condition in Study 4 before they performed the experiment (*M=60.80, SD=10.74*).

Participants’ trust in robots in the control condition in Study 4 before they performed the experiment (*M=63.66, SD=11.70*).