The University Of Sheffield.

Attitudes toward robots – fiction or reality?

Stanislava Ivanova Naneva

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Science

Department of Psychology

March 2021

Declaration

I, the author, confirm that the Thesis is my own work except where work that has formed part of jointly authored publications has been included. My contribution and that of the other authors has been explicitly indicated below. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not been previously presented for an award at this, or any other, university.

The systematic review presented in Chapter 2 has been published in a peer reviewed journal:

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

The work in Chapter 2 was jointly authored and is part of the second author's, Marina Sarda Gou, own thesis which has been submitted to The University of Sheffield. I (listed as the first author) carried out the primary data extraction, quality assessment, and data analysis as well as the writing up of Sections 2.2 and 2.3. Marina Sarda Gou (listed as the second author) carried out the literature search, all aspects of the screening, and the emailing of authors with requests for missing data and papers. Both authors contributed in roughly equal measure to the writing up of Sections 2.1 and 2.4. Research supervisors, Prof. Thomas Webb and Prof. Tony Prescott, are listed as co-authors who reviewed and provided comments on all aspects of the work and manuscript from conception to publication.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my primary supervisor, Professor Thomas L. Webb, for his guidance, patience, and unwavering support throughout the writing of this thesis. I could not have asked for a better supervisor and mentor during the past four years and will not forget all that I have learned under his care.

I would also like to offer my sincerest thanks to my second supervisor, Professor Tony J. Prescott, for his support, insightful advice, and incredible expertise. Without him, I would have not had the many opportunities to engage with the wonderful team at Sheffield Robotics, nor to experience the ground-breaking research happening there.

I would like to also thank my colleagues and peers at the Department of Psychology for their kindness and support. In particular, Marina Sarda Gou with whom I collaborated on a part of this thesis. My gratitude also goes out to the professional and academic staff at the Department of Psychology for their hard work in making a difficult Ph.D. journey that much easier for myself and others. I would also like to thank The University of Sheffield for making this thesis possible by funding my studies.

Finally, I would like to thank my family and Harry for their continuous support.

Summary

One challenge to understanding how people think and feel about robots is the wide variety of robotic systems and the general public's lack of direct contact with them, which leaves open the question of whether people's attitudes toward robots are shaped by exposure to real robotic systems or not. This thesis presents evidence that people not only internally represent the concept of robots in varied ways but also that such representations may account for some of the reported variability in people's attitudes toward robots. Chapter 2 presents a systematic review that demonstrated said variability by quantifying people's attitudes toward social robots and highlighted a number of factors that have arguably not been sufficiently explored in the literature. Namely, the influence of people's individual representation of robots and the potential impact of fictional and non-fictional depictions of robots on people's attitudes. Chapter 3 explored both of these factors by presenting a semantic network reflecting the social representation of robots that supported the diversity of individuals' representations and provided insights into the stable structure of the social representation which was divided into five distinct modules of meaning. These modules were further interrogated in the same chapter via the thematic analysis of semi-structured interviews which demonstrated both the role of fiction in people's individual representation of robots as well as the impact of said representations on people's attitudes. In order to investigate the impact of fictional robots on people's attitudes, three pilot studies that manipulated the perceived fictional status of identical robots through indirect contact were conducted and reported in Chapter 4. The methodology was then implemented in an experimental study that tested whether the perception of the fictionality of the robots had an effect on participants' attitudes (reported in Chapter 5). Findings showed that when robots are perceived as non-fictional, participants reported more positive explicit (but not implicit) attitudes toward those specific robots and toward robots in general than when the same robots were perceived as fictional. Chapter 5 also describes the findings of a second experimental study that primed participants with images of either fictional or non-fictional representations of robots. Findings supported the preference for non-fictional robots found in the previous study but in this case, there was a change in participants implicit, rather than explicit, attitudes. The implications of these findings are discussed in Chapter 6.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Chapter Summary

In Chapter 1, the current literature demonstrating the variability of people's attitudes toward robots has been briefly presented. The possible impact of fictional representations of robots on people's attitudes and the main research questions addressed in this thesis are also introduced.

1.1 Social Robots

Recent advances in animatronics, robotics, and artificial intelligence (AI) have seen a shift from an emphasis on the mechanical functionality of robots (e.g., manufacturing robots), to aspects that are more social (e.g., robots that can care for people and provide companionship). Robots with some ability to interact socially by learning and responding appropriately to social behaviours already exist (Duffy, 2003). Such robots are often referred to as *social robots*; definitions vary but generally any fully or semi-autonomous system that is capable of producing behaviour that is congruent with cues from its social environment could be defined as a social robot (Duffy, 2003). Currently, social robots are being designed and tested for use in customer service settings such as shops and public spaces. These robots are often humanoid in appearance and capable of interacting verbally with humans to aid social interaction (e.g., Pepper, Spencer). With advances in AI and further development of existing systems, social robots have the potential to alleviate human labour shortages in various sectors or provide an alternative to human labour.

The development of social robots can be valuable in not only enhancing customer service but also in providing health and caring for the elderly (Prescott et al., 2012). With an ageing population that has been estimated to keep increasing (Giannakouris, 2010), healthcare systems are now struggling to cope with providing care for the elderly and disabled. A report by the charity, Age UK (2017), indicated that the number of elderly individuals who were receiving social care support funded by local authorities decreased by 6.1% between 2005 and 2014. The report also indicated that this decline was likely to be the result of increasing numbers of elderly individuals, government funding cuts (relative to the demand for care), and less spending on elderly care by local authorities (Age UK, 2017). Social robots could potentially be one way to address the welfare and economic issues that arise with the increase in healthcare demands. Although the technology is not yet ready to be implemented on a large scale and there is a multitude of practical, legal, and ethical issues yet to be resolved, small-scale field tests of social robots in care settings are already being carried out (Broadbent et al., 2010, 2012; Jayawardena et al., 2012). Although social robots have the potential to relieve carers' responsibilities and increase the autonomy of elderly and disabled individuals (Prescott et al., 2012), there are many concerns surrounding the design of such robots and the extent to which they will be accepted by prospective users. Many people have reservations about the application of robotics in healthcare settings (and other socially focused sectors) due to potential job loss, loss of autonomy, privacy, and other ethically challenging issues (European Commission, 2012). While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (for example, they may think that robots are considerably more advanced than what is currently possible). This can lead to low levels of trust, acceptance, and negative attitudes toward robots, as well as to poor interaction between people and robotic systems in healthcare contexts and beyond (Broadbent et al., 2009).

If robots are to become increasingly present in our daily lives - and even take on some of the roles humans hold in society - then there is a need to understand people's attitudes toward robots and the impact that these attitudes have on the acceptance and implementation of robotic systems. Studies exploring the long-term effect of social robots in healthcare and educational settings have also become available. In a review of such longitudinal studies, Leite et al. (2013) found tentative evidence for the acceptance and positive impact of social robots in various therapeutic and educational settings. The past few decades have seen a growth in research (that has been reviewed in Chapter 2) focusing on social human-robot interaction, public opinion of the use of robotics, and a more user-centred approach to designing social robots. However, there are still research avenues that have not been fully explored in regards to the factors that influence and shape people's attitudes toward robots.

1.2 Variability of Attitudes Toward Robots

Research into people's attitudes toward robots reveals a somewhat ambiguous picture and it is difficult to say whether people, in general, have a largely negative or positive view of robots. As detailed below, attitudes toward robots likely depend on the intended use of the technology in question and a number of other factors that have been addressed in Chapter 2.

A now somewhat dated survey by Takayama et al. (2008) found that the majority of people favoured the use of robots for jobs which required the retention of large amounts of information, repetitive and perceptual based work, or service based jobs such as ushering. However, for jobs that required adaptive thinking, problem solving, or advanced social abilities such as diplomacy, robots were much less favoured. In fact, people preferred that robots be utilised for jobs that require memorisation or as a support to their own work, rather than as a replacement for humans. Similarly, according to a survey by the European Commission (2012), 60% of EU citizens believe that the use of robots in the care of children, elderly people, and people with disabilities should be banned. The potential use of robots in education and healthcare was also somewhat unpopular with 34% and 27% of respondents stating that it should not be allowed. On the other hand, the application of robotics in space exploration and manufacturing was met with wider support by half of those who took part; as was using robots in military and security activities, as well as search and rescue activities (identified as a priority by 41% of surveyed EU citizens). In line with these figures, less than half of the respondents stated that they would feel "totally comfortable" accepting assistance from a robot at work, whilst 86% stated that they would feel "totally uncomfortable" with having their children or elderly parents minded by a robot. Similar attitudes toward the use of robots for specific roles have been noted by other studies. For example, Enz et al. (2011) found that attitudes toward the future social roles that robots may hold were negative when participants were presented with scenarios that implied that in the future robots would be as equal in capabilities or social standing as humans (e.g., legally recognised as citizens). Work alongside humans was again noted as preferable to humans being completely replaced in certain social roles. Interestingly, attitudes toward robots in care roles such as nursing were relatively positive but participants were still more supportive of the use of robots for dangerous and menial tasks such as manufacturing. One possible explanation for this difference is that Enz et al.'s sample contained a larger proportion of professionals in the fields of nursing, medicine, and technology when compared to the study conducted by the European Commission. Considering that some scenarios described robots as helpers in tasks such as assisting people with personal hygiene, a job which usually falls on nursing staff, it is possible that the role of robots in fields relating to healthcare were perceived as a chance to alleviate some of the pressures in those fields of work.

A recent longitudinal study by Gnambs and Appel (2019) indicated that there was a general decline in favourable attitudes toward robots among EU citizens between 2012 and 2017. This trend was particularly notable for attitudes toward the use of robotic systems at work (e.g., manufacturing), indicating a growing scepticism surrounding the use of robots to alleviate human labour shortages and assist people performing challenging manual labour. Despite this decline, attitudes toward the use of robots in health and elderly care sectors were still by far the most unfavourable in general. Based on existing research, the application of robots in select domains that tend to receive the largest support from people are also the ones in which robotic technology is already present to a large degree (e.g., manufacturing). Similarly, the domains of application that receive the least support are ones in which robots are not well established, if at all. For example, while the development of robots for education, elderly care, social assistance, and healthcare is underway, they are by no means commonplace. As such, there might be hidden difficulties with trying to gain insight into attitudes toward technology that already exists and is more likely to be familiar to the general public versus technology that is not only uncommon, but also not yet a practical reality. In other words, attitudes toward the application of robots in domains such as elderly care and domestic service is typically, but not necessarily, prospective in nature.

People's beliefs about what the future may hold are likely to be varied and not necessarily grounded in the reality of the currently available technology and its application. To further complicate matters, what roboticists and researchers may consider to be suitable avenues for the development and use of social robots may not necessarily align with public opinion. This presents somewhat of a dilemma as it could lead to the creation of robotic systems that are either not suitable for their intended use or completely unwanted by some or all of the intended end-users. In recent years, user-centred approaches to the design of robots have become increasingly popular and are intended as a means of avoiding the above mentioned problems. Such approaches generally acknowledge that the 'experts' (i.e., roboticists and / or researchers) may not necessarily know what the optimal design for a social robot is for a specific group of stakeholders in a specific context. In essence, robot development is shifted from a top-down, expert-first approach to a more collaborative, bottom-up approach. It is generally accepted that involving stakeholders in the design process yields a better match between the needs of the users and how the technology addresses those needs. Focus groups, individual interviews, and field experiments are just some of the methods that have been utilised to inform the design and function of robotic systems via a user-centred approach (Salvini et al., 2010; Cavallo et al., 2013; Lehmann et al., 2013; Reich-Stiebert, Eyssel, & Hohnemann, 2019). In a small-scale study looking at the preferences of elderly users and roboticists for the design and function of companion robots, Bradwell et al. (2019) observed a marked difference between the two groups. These differences extended beyond simple preferences for features (e.g., fur or no fur) and into attitudes toward specific functions such as speech. The focus groups conducted by Bradwell et al. suggested that the observed differences in preference of zoomorphic companion robots between end-users and roboticists may stem from a difference in the two group's underlying beliefs, assumptions, and experiences. For example, from a design perspective, speech is not necessary for a robot to interact socially with the user and may be difficult to execute in some contexts. However, the users' attitudes toward robots capable of speech was linked to their experiences of loneliness as elderly people and the act of relieving loneliness via communication which could be one of the needs addressed by companion robots. Given that the elderly users in this case expressed surprise at the lack of speech capabilities of the robots suggests that their expectations and understanding of robots is not necessarily congruent with that of researchers. Since the general public are unlikely to come into contact with more advanced robotic systems (social robots in particular) due to their lack of availability, fictional and media representations of such robots could be particularly influential in attitude formation as most people are more likely to be familiar with fictional rather than real robots. More specifically, fictional portrayals of robots may shape the way individuals internally represent robots as a broad category.

1.3 Attitude Representation Theory

Lord and Lepper's (1999) Attitude Representation Theory could help to explain variability in attitudes toward robots in recent research (Nomura et al., 2005; Takayama et al., 2008; Enz et al., 2011; European Commission, 2012) as it suggests that the way that people represent the attitude object (here, a robot, or robotics in general) likely influences their reported attitudes toward that object.

Attitude Representation Theory is based on two key postulates – (i) *representation* and (ii) *matching*. According to Lord and Lepper, an individual's response to (or evaluation of) a stimulus (e.g., a robot) will depend on not only the apparent characteristics (e.g., a humanoid appearance) and context (e.g., technological exhibition) of that stimulus at that particular time, but also on the representation that the individual already has of the stimulus or respective category. Lord and Lepper further argued that the representation is also relevant when individuals are asked to make broader evaluations of the category to which a stimulus may belong, because they may rely on a specific and subjective representation. For example, if someone is asked to give their opinion about robots (i.e., a broad category) they may base their response on a representation of that category (i.e., an exemplar) that they have generated (e.g., Terminator). The potential for a wide variety of representations of robots to come to mind may explain why attitudes toward robots have been demonstrated to be highly variable between individuals and across the various sub-categories of robots (Takayama et al., 2008; Enz et al., 2011; European Commission, 2012; Pino et al., 2015).

The second postulate of attitude representation theory is that of *matching*, which suggests that the consistency of individuals' evaluations (or responses) to a category depend on how closely the relevant stimulus and the generated exemplar at one instance match the stimulus and exemplar at a different instance (Lord & Lepper, 1999). That is, as long as the stimuli are similar and the exemplars retrieved at different instances are the same (i.e., there is 'matching'), then the resulting attitudes toward the category to which the stimuli belong are likely to remain consistent. In contrast, the variability in attitudes toward robots could potentially be the result of different exemplars of robots being brought to mind at different times and in different contexts (i.e., a lack of matching). A lack of matching seems particularly likely considering the relatively wide variety of robots and contexts in which a person might encounter them. For example, a person could encounter an industrial robotic arm at their place of work and later see a photograph of a humanoid robot in a newspaper. It is therefore critical to understand what representations people bring to mind when asked about robots and how these shape their resulting beliefs and behaviours toward such technology. One way to interrogate how representations are formed is by considering the role of fictional depictions of robots.

1.4 Robots in Fiction

Fictional material can provide an opportunity to discuss various ethical and practical issues; even predicting or inspiring technological developments. For example, the communicators depicted in the science fiction TV series, Star Trek, inspired the design of the first mobile phone (Cuneo, 2011). Considerable fictional material has been produced portraying the possible futures associated with a widespread application of robotics. Often, said future is either envisioned as a robot-dystopia (e.g., Matrix, 1999) or as prosperous world maintained by subservient robots (e.g., I, Robot, 2004). The way robots are depicted in fiction can vary greatly, from almost indistinguishable in appearance and intelligence from humans (e.g., Ava in Ex Machina, 2014) to swarms of maleficent squid-like robots (e.g., the Sentinels in Matrix, 1999).

Although the way robots are portrayed in fiction varies and so do the narratives within which they are depicted, there is evidence of some consistency between portrayals in science fiction films. Kriz et al. (2010) conducted a content analysis of 12 science fiction films and found that, whilst the majority of main character robots were portrayed as having good or superior-to-human cognitive abilities (e.g., problem solving), many lacked human-like social behaviours (e.g., conformity). Kriz et al. go on to suggest that familiarity with fictional robots that are consistently represented in a certain way, may influence people’s expectations about real robots’ cognitive and social abilities. To investigate this idea, Kriz et al. showed seventy-seven engineering students the PeopleBot (a semi-humanoid robot) and asked them to rate how likely they thought the robot was to exhibit a number of social and cognitive abilities (e.g., “How likely is it that this robot could understand that it is a robot?”). Participants’ responses followed the same patterns as the characteristics that fictional robots typically portrayed (i.e., PeopleBot more likely to have cognitive rather than social abilities). It should be noted that only a small selection of science fiction films was included in the study and the majority of films (n = 8) were produced prior to the 21st century. Whilst no formal investigation of the differences in fictional representations of robots over time exists, it is quite possible that the way robots are portrayed has changed over time. If this is indeed the case, it is also likely that people from different age groups may be more familiar with some robot depictions than others. However, whether such a difference results in any variability in people’s attitudes toward robots is unclear.

1.5 Can Fiction Influence People's Attitudes Toward Robots?

Some evidence of differences between people's perceptions of fictional and non-fictional robots already exists. For example, DiSalvo et al. (2002) found that, on average, the heads of fictional robots were rated as more human-like than those of non-fictional robots. This, to an extent, is not particularly surprising as fiction is largely free of the physical constraints that real technology is under. In other words, creating robots is much easier with the help of digital and cinematic effects. Although DiSalvo et al.'s aim was to inform the design of real robots, their findings also suggest that there is a difference in how people perceive fictional robots as compared to non-fictional ones. This is important since anthropomorphism (i.e., how humanlike a robot is) has been shown to influence people's attitudes toward robots (Duffy, 2003) and if fictional robots do indeed influence people's attitudes, this finding may be worth keeping in mind.

Video material in the public domain has already been used to investigate how fictional and non-fictional robots are perceived. For example, Mubin et al. (2015) performed content analysis of comments left on YouTube videos that depicted two fictional (HAL9000 and Astro Boy) and two non-fictional (Nao and Shakey) robots. For each category, they selected the most and least anthropomorphic robots (as guided by the taxonomy presented in Zhang et al., 2008). Comments were coded into levels of three main categories: (i) the topic/content, (ii) the degree of anthropomorphism, and (iii) the valence (positive or negative). Contrary to their initial hypothesis, Mubin et al. found that the videos of the two non-fictional robots (Nao and Shakey) generated more engagement and positive interest compared to the videos of the fictional robots (HAL900 and Astro Boy). This may be for a variety of reasons. It could be that fictional characters (including robots) may be more prone to negative comments due to their affiliation with particular sci-fi movies and their role within them. Alternatively, it could be that nonfictional robots are less familiar to viewers and, as such, are more novel, thus leading to more engagement with videos depicting non-fictional robots. Sadly, no information can be extrapolated about the commentators' general attitudes toward robotics, which could be important to understanding the way that they perceive fictional and nonfictional robots.

Evidence from Riek et al. (2011) suggests that science fiction could have an impact on people's attitudes toward robots. In a survey of 287 people, Riek et al. found that participants who reported having watched a larger number of films about robots also reported more positive attitudes toward robots (as indicated by low NARS scores). It should be noted that this correlation was weak but nonetheless indicates a possibility that viewing fictional portrayals of robots could play a role in shaping people's attitudes. Of course, since this was a correlational study, it is also likely that people who already hold positive attitudes toward robots are more likely to select and watch robot themed films. Riek et al. also point out that they only asked participants to indicate if they have watched a specific film and made no record of the number of times a film was watched, how long ago it was viewed, or any other potentially confounding information.

Since the majority of end-users are unlikely to have direct experience with the types of robots being developed, it is interesting to see whether their opinions are guided by experiences with fictional depictions of robots. Findings from studies using user-centred approaches to robot design may help to answer this question. For example, focus groups with elderly people conducted by Wu, Fassert, and Rigaud (2012) indicate that participants' attitudes toward the use of specific robots (fictional and real) for care of the elderly was influenced by the appearance of the robot. Despite seeing some robots for the first time, participants showed consistent behavioural responses (e.g., smiling and laughing) towards robots with specific characteristics (e.g., those with a "cute" appearance). Whether participants' attitudes and responses were mediated by fictional / media representation of robots is unclear and as such does not provide much insight into the impact of fiction on attitudes toward robots. However, the findings do suggest that participants made assumptions about the capabilities of the robots that they saw and were confused by the apparent gap between some of the robots' (sophisticated) appearance and their (less sophisticated) function. This implies that, even if participants had no direct experience with or knowledge of the robots that they saw, they still associated pictorial representations of robots (i.e., representations reflecting their appearance) with specific positive or negative evaluations. Such evaluations are likely to play a role in the acceptance of assistive technology by elderly users and impact the intention to use robots in various settings such as healthcare (Louie et al., 2014; Smarr et al., 2012).

A further investigation of the impact of science fiction and mass media on the user-centred design of robots was conducted by Bruckenberger et al. (2013) via a focus group, interviews, and an online survey. They were particularly interested in the impact of "good" (i.e., robots as heroic and helpful) and "bad" (i.e., robots as evil and antagonistic) depictions of robots in science fiction. Bruckenberger et al. found that either type of depiction lead to ambiguous or double-minded attitudes toward robots that were not based on reality. Findings suggested that because depictions in science fiction generally represent robots as human-like, they can give rise to anxieties about the role of robots as either villains who can harm humanity or as competent agents that can replace humans. Conversely, participants who reported more familiarity with real robots represented by the mass media generally had a more grounded view of robotics and decreased anxiety toward robots. However, the majority of participants were not familiar with real robots as depicted in mass media. Furthermore, participants who reported that they had encountered media coverage of real robots did not generally accept those representations as prototypical of the robot category. As fictional representations of robots are encountered more often than real robots or non-fictional media and are therefore more familiar, they may also be more likely to be accepted as an exemplar of the robot category regardless of how reflective of reality they are. If this is true, Bruckenberger et al.'s findings can have implications for the representation postulate of the Attitude Representation Theory.

1.6 Research Questions

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of robots play a significant role in shaping people's attitudes, acceptance, and expectations of such technology. This can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology (Kriz et al., 2010). While many studies have looked at people's attitudes toward robots in different contexts (see Chapter 2), it is still unclear what role fictional representations of robots play in the formation and measurement of those attitudes.

Therefore, this thesis aimed to investigate the nature of the representations that people have of robots, their origin, and how such representations influence people's attitudes toward robots. Specifically, this thesis aimed to answer the following questions:

1. What representations typically come to mind when people think about robots?
2. Do those representations influence people's beliefs about, and attitudes toward, robots?
3. Do fictional portrayals of robots influence how people represent robots and people's beliefs about, and attitudes toward, robots?

Chapter 2: Study 1 - A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots

Chapter Summary

Chapter 2 presents a systematic review of 97 empirical studies that quantified people's attitudes toward social robots and explored various moderating factors known to affect attitudes. The findings suggest that although people have overall positive attitudes toward social robots, there is considerable variability in their attitudes that cannot be entirely explained by factors such as the robots' domain of application and design, the type of exposure to the robot, and the individual characteristics of participants. Although the work presented in Chapter 2 resulted in a broad overview of the existing research on people's attitudes toward robots, it did not lead to any additional insights on how, and if, the way people internally represent the concept of robots affects their attitudes.

2.1 Introduction

According to a widely-reported large-scale survey (European Commission, 2012), a substantial proportion of EU citizens have negative attitudes toward the use of robots within healthcare and other fields that are traditionally dominated by humans. There have also been suggestions of a growing anxiety among the public that automation, enabled by robotics, will lead to a significant loss of jobs (Ebel, 1986; Broadbent et al., 2009). As we will explore in this review, attitudes toward robots appear mixed, likely depend on the setting and question asked, and in some cases are 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 toward robots should therefore help to inform future research, development, and deployment of robotics in various domains of public and private life.

This review focused on social robots, due to their increasing use in various settings such as healthcare, entertainment, and customer service (Takeda et al., 2007; Hancock et al., 2011; Pieska et al., 2013). 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 (Nomura et al., 2006; Ray et al., 2008). Nevertheless, social robots garner attention from the media and general public alike, and have sparked debate about their potential impact on society (Zhao & Yi, 2006; Nørskov, 2017). We have defined 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 Appendix A). 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, 2010), and reflect the same broad construct (Li et al., 2010; Gaudiello et al., 2016; Gombolay et al., 2018; Herse et al., 2018), which is people's perception or evaluation of robots.

2.1.1 Attitudes toward 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 et al., 2011). At the same time, some studies have found positive attitudes toward robots performing jobs that demand more social skills (Enz et al., 2011; European Commission, 2012). 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). Differentiating 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).

2.1.2 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., 2006; 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., 2006) 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 et al. (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.

2.1.3 Trust in social robots

Trust has also 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 et al., 2015). Trust is likely to be particularly important in relation to social robots, especially in healthcare, where trust has been associated with patient satisfaction and therapeutic effectiveness (Hall et al., 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., 2011). However, the impact of trust in relation to social robots specifically has not been reviewed.

2.1.4 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; Venkatesh et al., 2003; Heerink et al., 2010). 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 et al., 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.

2.1.5 What factors influence people's attitudes toward robots?

Several factors are likely to be associated with people's attitudes toward, trust in, acceptance of, and anxiety toward 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, or healthcare), and the design of the robot (e.g., humanoid or anthropomorphic). We expand on these potential factors below.

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):

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 toward social robots in general; de Graaf et al., 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 et al., 2015).

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 et al., 2017, Savela et al., 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; 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 et al., 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 et al., 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 et al., 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 et al., 2016).

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., 2011; 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 et al., 2014).

Anthropomorphic - a robot that imitates some parts of the human body and can be subject to anthropomorphisation by the user (e.g., a robot with a human-like face; Dunst et al., 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., 2007).

Geographical location. The cultural background and nationality of users may contribute to the variability in people's attitudes toward (Bartneck et al., 2005), trust in (Li et al., 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).

Sample characteristics. Attitudes towards robots also likely vary according to demographic factors such as users' age and gender (de Graaf & Allouch, 2013). For example, men generally tend to have more positive attitudes towards robots than women (May et al., 2017). Similarly, young adults tend to have more positive attitudes toward 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 (Miller et al., 2012) which is why the present review also attempted to investigate this factor.

2.1.6 Aims of the review

This review has expanded on earlier efforts to understand people's beliefs about social robots (e.g. Broadbent et al., 2014; Chen & Chan, 2011; Hancock et al., 2011; Savela et al., 2018) 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.[1] 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.

2.2 Method

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

2.2.1 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

[1] A traditional meta-analysis was not viable since traditional effect size metrics cannot be used to describe the average level of a given variable (e.g., the valence of people's attitudes); only the extent to which it is influenced by a manipulation (e.g., effect size Cohen's *d*) or is related to another measure (e.g., effect size *r*).

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 to and managed via EndNote where duplicates were removed prior to screening the research articles. Figure 2.1 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.

2.2.2 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.

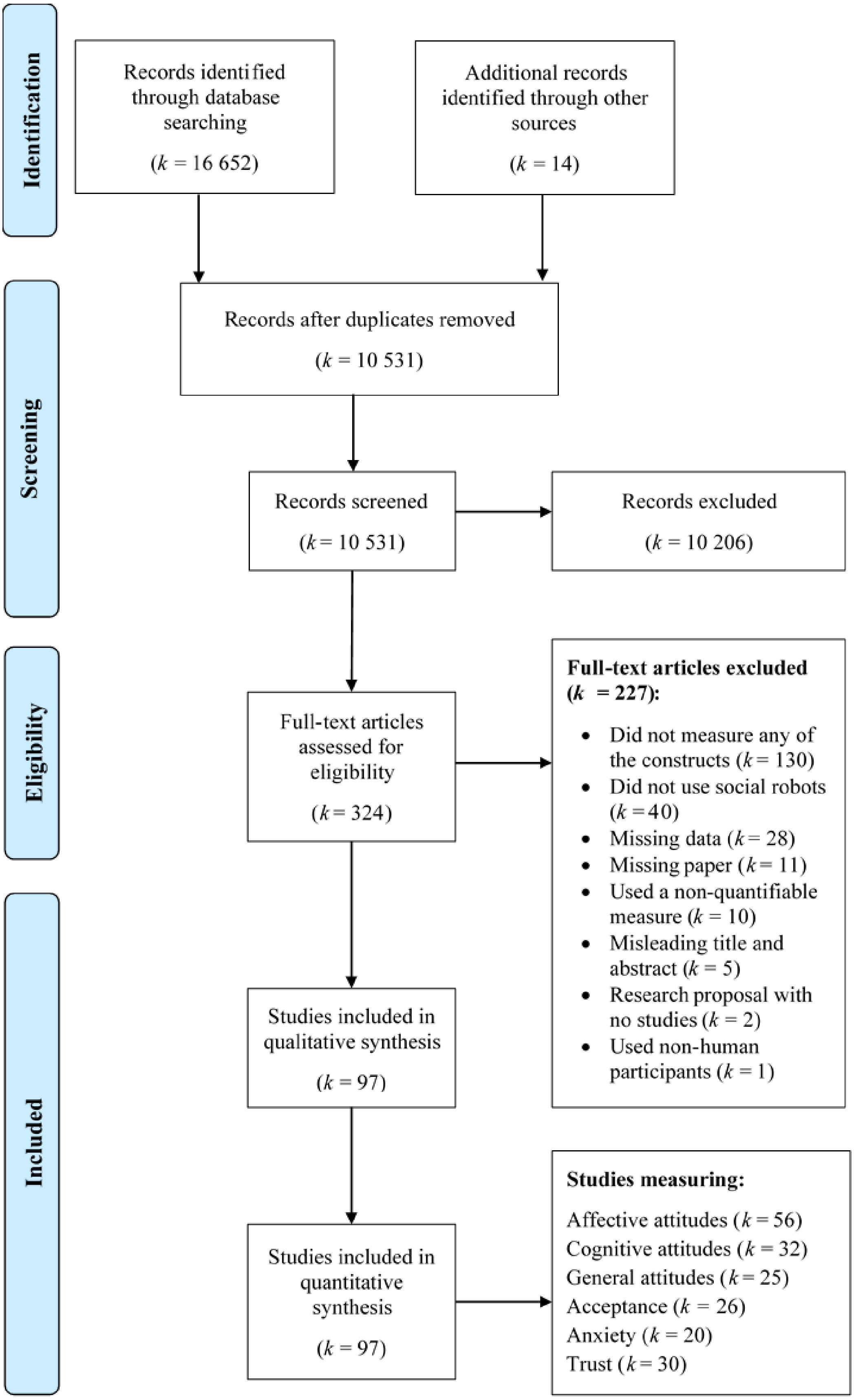
We first extracted bibliographic information from the articles, this included the date of publication, the country where the research was conducted, the sample size and demographics of the sample (i.e., mean age, gender, and cultural or ethnic background), the domain of application, the design of the 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.[2]

The methodological quality of the primary studies (i.e., risk of bias) was assessed using the tool described in Appendix B. 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.

[2] Measures of acceptance typically overlapped with behavioural attitudes (i.e., reflected people's observable or self-reported behaviours and / or intentions to use robots). Therefore, behavioural attitudes were omitted as an outcome and instead we examined evidence on the acceptance of social robots, which was easier to identify.

Figure 2.1

PRISMA diagram showing the flow of studies through the review



2.2.3 Calculating and interpreting rescaled and "standardised" outcomes

The present review sought to quantify the valence of people's attitudes toward robots and compare this between different contexts, methods, and samples. As such, 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 and $\bar{x}$ and s denote the sample mean and standard deviation extracted from each study.[3]

---

[3] Note that *MR* is the numerical value for each scale that indicates a neutral attitude. If a scale measured negative attitudes, then $\bar{x}$ was reversed prior to calculating $\overline{x_s}$ by adding the maximum (x_{max}) and minimum (x_{min}) possible values of each scale and taking away the $\bar{x}$ (e.g., for a 1 to 5 scale with a mean of 2, the reversed score would be (1 + 5) - 2 = 4.

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

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

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.

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

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

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

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

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

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.

2.2.4 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 $p < .05$.

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

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

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

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

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

Table 2.1 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.

2.2.5 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 Appendix A) 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 2.1.

2.3 Results

2.3.1 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.

2.3.2 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., 2006) 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 toward social robots in general (e.g., Dinet & Vivian, 2014) or specific types of social robots (e.g., domestic robots; de Graaf et al., 2016), while the rest measured participants' attitudes toward specific social robots (e.g., NAO; Torta et al., 2014).

Figure 2.2

Plot of pseudo-standardised means ($\bar{x}_s$) 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 ($\bar{x}_w$) for affective attitudes and the error bars represent 95% $CI_{\bar{x}_w}$

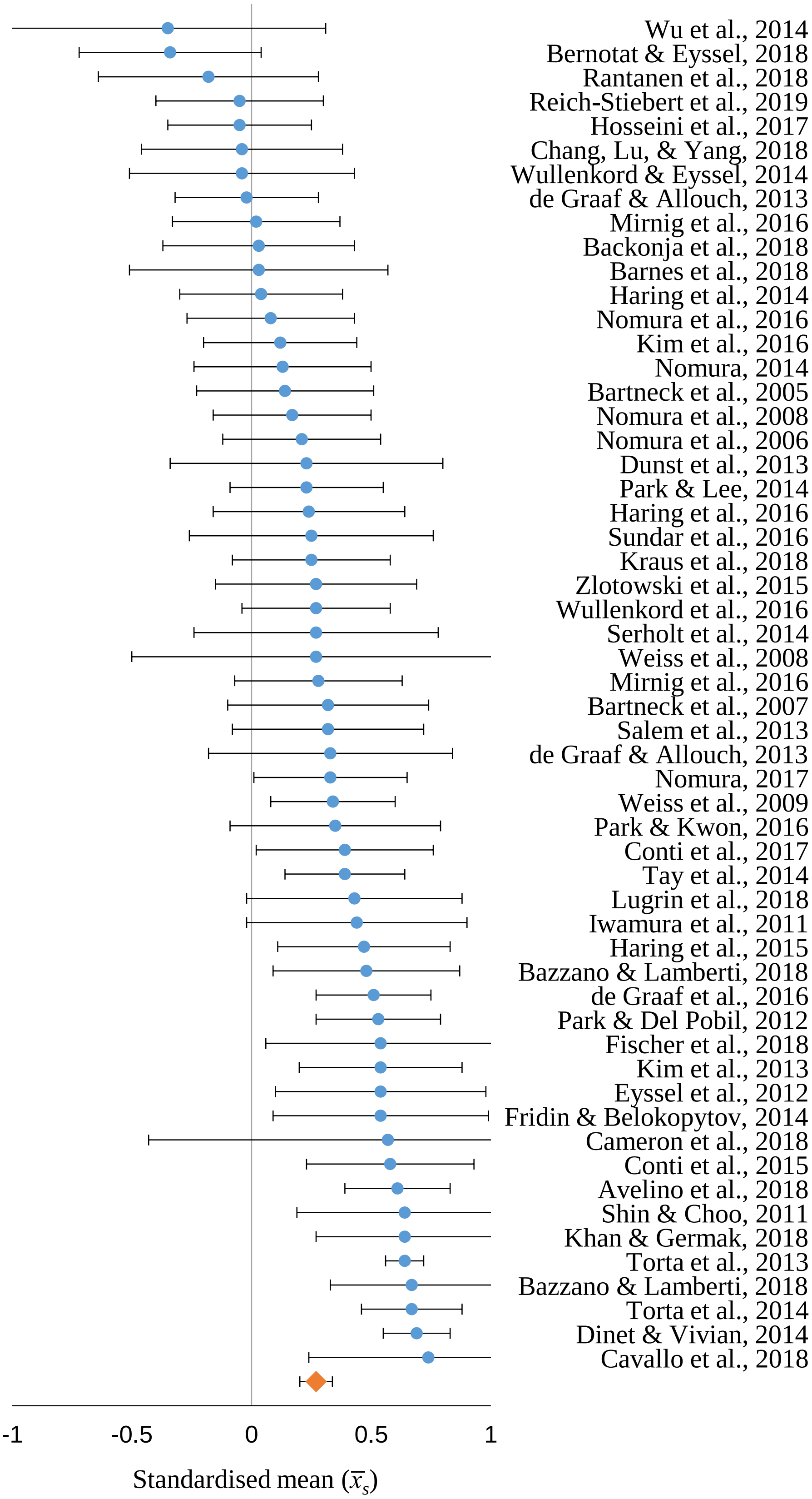
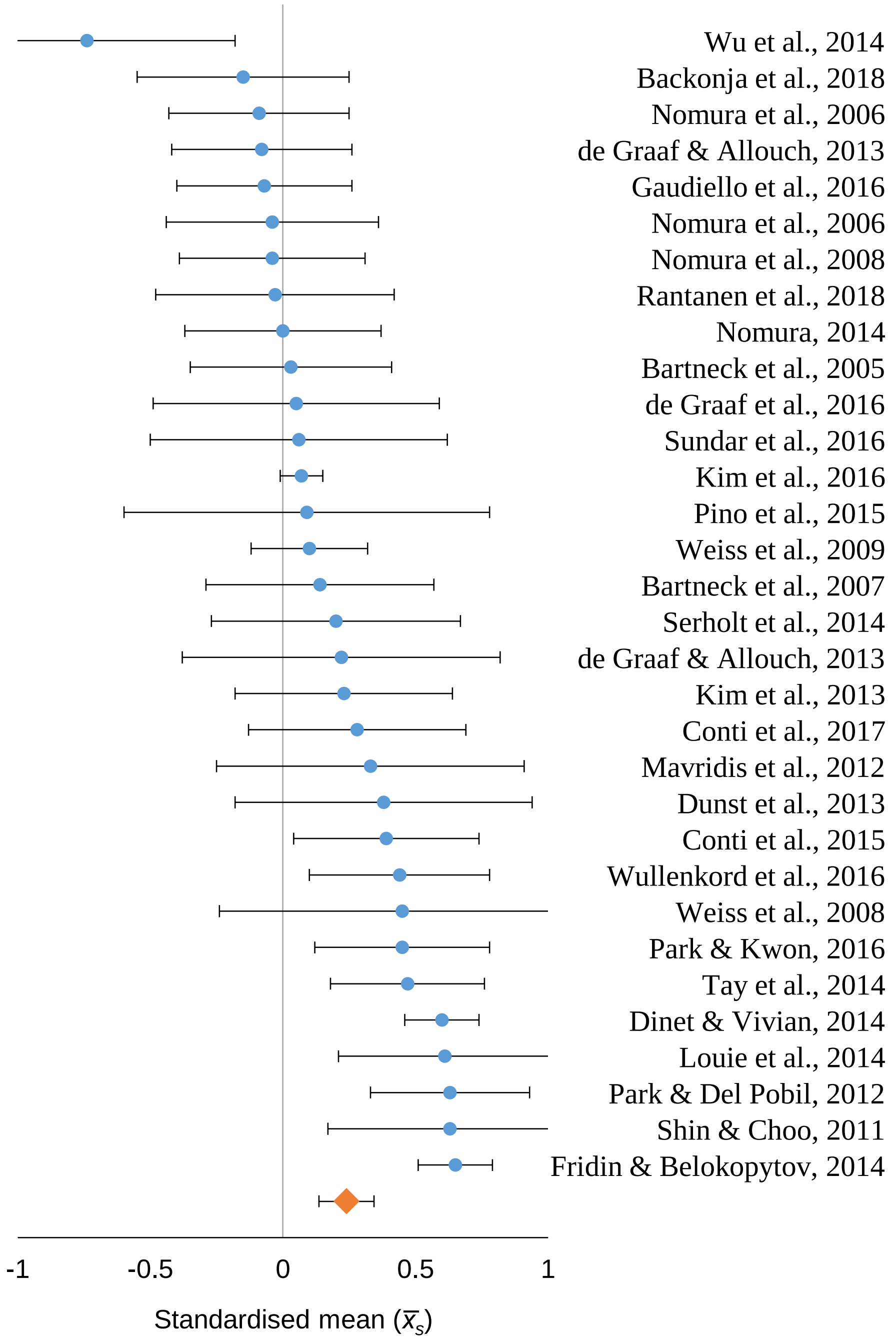


Figure 2.3

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 95% $CI_{\bar{x}_w}$



The average weighted mean for affective attitudes was $\overline{x}_w = 0.27$ (see Figure 2.2), 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., $\overline{x}_w < 0$) and only 16 studies (29%) had a mean of $\overline{x}_s > \pm 0.50$, signifying that people held substantially positive or negative affective attitudes.

2.3.3 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 (Shin & Choo, 2011; Tay et al., 2014; Conti et al., 2017).

The average weighted mean for cognitive attitudes was $\overline{x}_w = 0.24$, indicating that, in general, people had slightly positive cognitive evaluations about social robots and their use (see Figure 2.3). The majority of studies (72%) found evidence for positive cognitive attitudes with one study by Nomura (2014), providing evidence for neutral cognitive attitudes ($\overline{x}_s \approx 0$).

2.3.4 General attitudes

Twenty-five studies (26%) measured attitudes toward 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 three studies (12%) that used the Implicit Association Test (IAT). The aggregated data (see Figure 2.4) indicated an average weighted mean of $\overline{x}_w = 0.07$, which suggests that people's general attitudes toward social robots tended to be neutral (bordering on slightly positive). Thirteen studies (55%) provided evidence of positive general attitudes (i.e., $\overline{x}_w > 0$) toward social robots while the rest provided evidence for negative attitudes, with one study reporting neutral attitudes (i.e., $\overline{x}_w = 0$).

2.3.5 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 ($\overline{x}_w$ = 0.24) indicated that, in general, people accept social robots but only slightly so. However, acceptance of social robots varied considerably (see Figure 2.5) and 42% of studies suggested that people did not accept robots (i.e., $\overline{x}_w < 0$). 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 ($\overline{x}_w$ = -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 dentified as a factor that impacts participants' intention to use robots (Venkatesh et al., 2003).

Figure 2.4

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 95% $CI_{\bar{x}_w}$

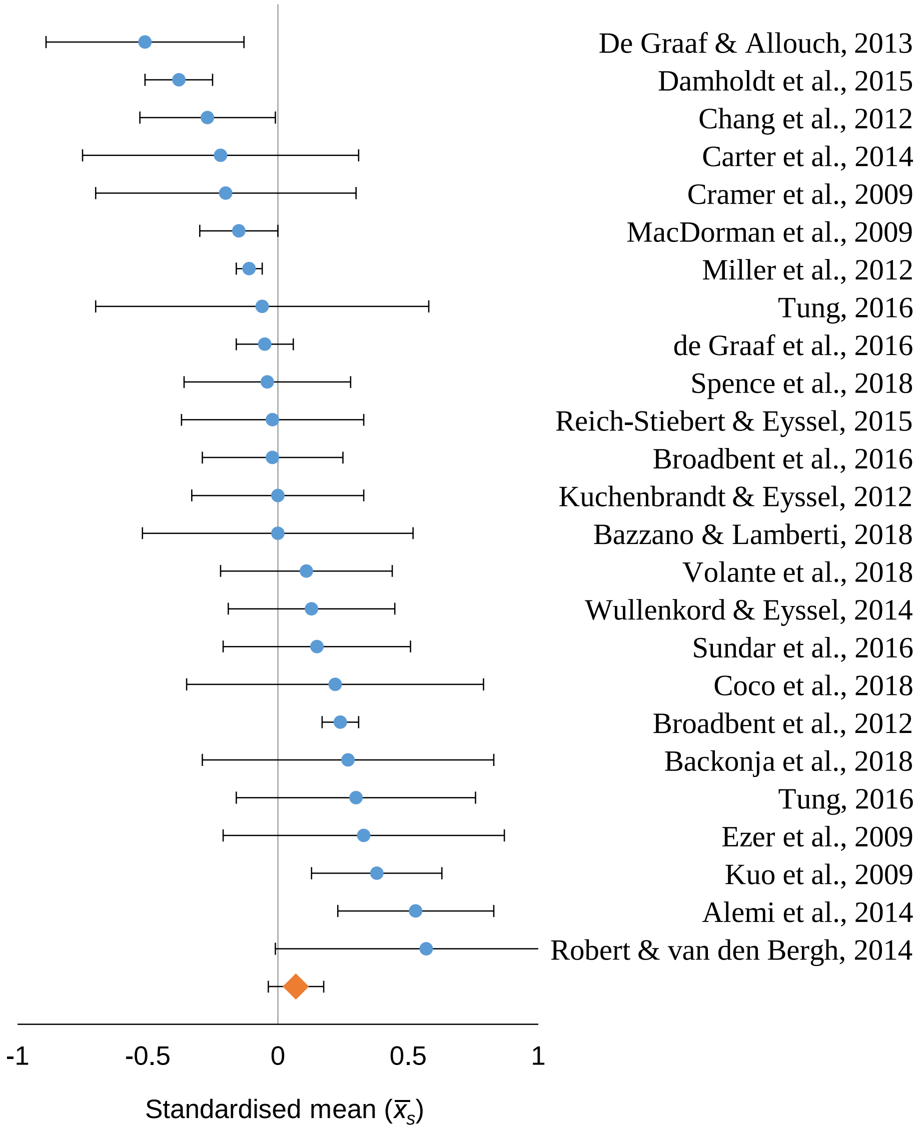
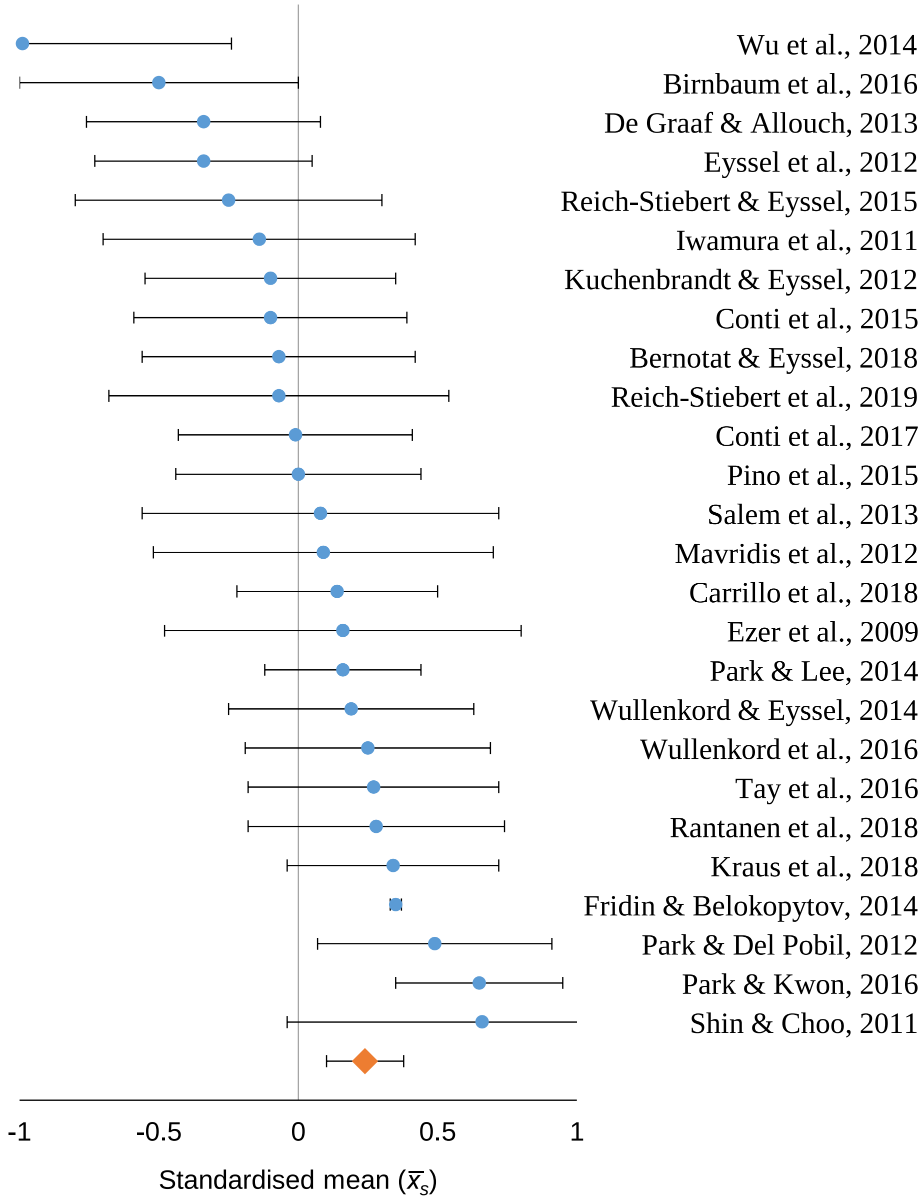


Figure 2.5

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 95% $CI_{\bar{x}_w}$

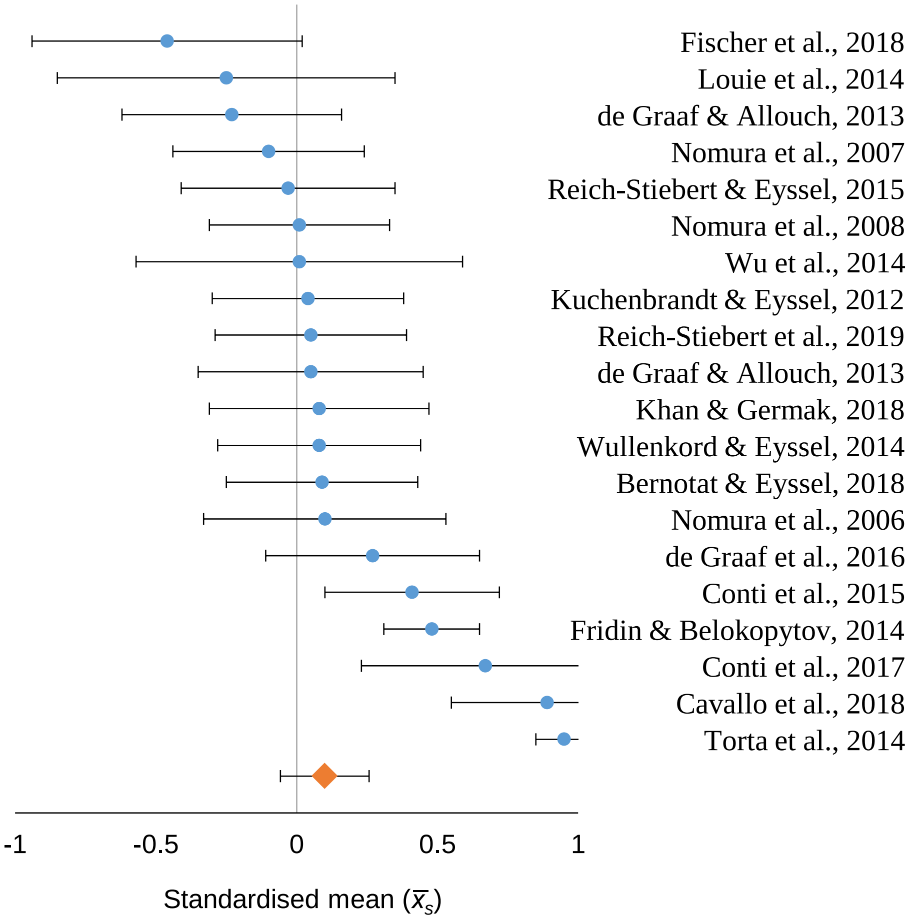


2.3.6 Anxiety

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., 2006) with ten studies (50%) having used some variation of the measure (Kuchenbrandt & Eyssel, 2012; Wullenkord & Eyssel, 2014; de Graaf et al., 2016). 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., $\bar{x}_w$ < ±0.10, see Figure 2.6). 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 2.6).

Figure 2.6

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 95% $CI_{\bar{x}_w}$



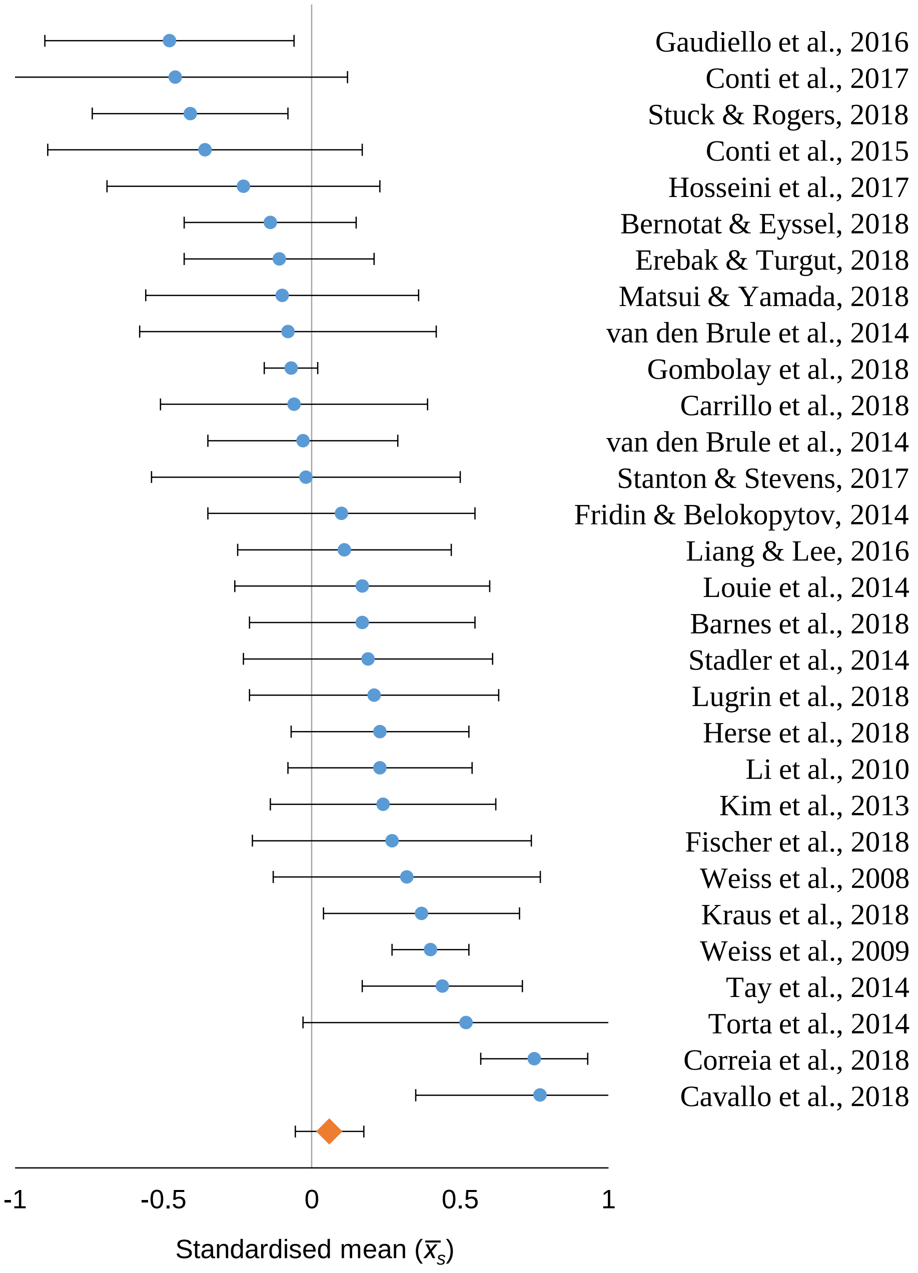
2.3.7 Trust

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, $\overline{x}_w = 0.06$, suggesting that, in general, people did not particularly trust or distrust social robots. However, the plot of all included studies (see Figure 2.7) indicated variation within and between studies with 43% of studies presenting evidence that people did not trust social robots (i.e., $\overline{x}_w < 0$).

Figure 2.7

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 95% $CI_{\bar{x}_w}$



2.3.8 Factors that influence the main outcomes

Table 2.1 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 Appendix C.

Table 2.1

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	0.40	0.08	4544	13	[0.35, 0.44]	0.35	0.09	4535	11	[0.29, 0.41]	-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	0.34	0.05	1807	26	[0.32, 0.36]	-0.13	0.08	1192	12	[-0.19, -0.08]	-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	-
Geographical location															
Australia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

France	0.58	0.12	232	3	[0.28, 0.88]	0.35	0.11	313	5	[0.21, 0.48]	-	-	-	-	-
Germany	0.22	0.12	668	10	[0.13, 0.30]	-	-	145	2	-	0.04	0.19	532	3	[-0.44, 0.52]
Italy	0.57	0.17	156	5	[0.36, 0.79]	-	-	80	1	-	-	-	18	1	-
Japan	0.21	0.13	1613	8	[0.10, 0.31]	0.05	0.07	1331	5	[-0.04, 0.14]	-	-	-	-	-
Netherlands	0.31	0.18	327	3	[-0.13, 0.74]	0.01	0.26	327	3	[-0.64, 0.65]	-0.09	0.10	386	3	[-0.35, 0.17]
New Zealand	-	-	-	-	-	-	-	-	-	-	0.23	0.07	220	3	[0.07, 0.40]
South Korea	0.43	0.21	350	3	[-0.08, 0.95]	-	-	270	2	-	-	-	-	-	-
Taiwan	-	-	226	1	-	-	-	-	-	-	-	-	578	2	-
USA	0.05	0.19	979	5	[-0.19, 0.29]	0.04	0.28	961	3	[-0.66, 0.74]	-0.10	0.04	2400	10	[-0.13, -0.07]

Table 2.1 (continued)

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

	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]	0.10	0.20	933	4	[-0.22, 0.41]	-	-	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]	0.65	0.08	314	9	[0.59, 0.71]	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	-
Geographical location															
Australia	-	-	8	1	-	-	-	-	-	-	0.11	0.23	108	3	[-0.45, 0.67]

France	-	-	36	2	-	-	-	11	1	-	-	-	60	2	-
Germany	0.03	0.16	967	8	[-0.11, 0.17]	0.04	0.18	712	4	[-0.24, 0.32]	0.31	0.22	106	3	[-0.25, 0.87]
Italy	-	-	80	1	-	0.57	0.22	123	3	[0.02, 1.12]	-	-	103	2	-
Japan	-	-	24	1	-	-0.01	0.20	376	3	[-0.52, 0.50]	-	-	87	1	-
Netherlands	-	-	60	1	-	0.03	0.23	327	3	[-0.53, 0.59]	-	-	237	2	-
New Zealand	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
South Korea	-	-	290	2	-	-	-	-	-	-	-	-	60	1	-
Taiwan	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
USA	-	-	177	1	-	-	-	-	-	-	-0.08	0.08	130	5	[-0.18, 0.03]

2.3.8.1 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 ($\bar{x}_m$ = 0.22) than in studies with direct ($\bar{x}_m$ = -0.14) or no HRI ($\bar{x}_m$ = -0.10). This lack of overlap between the $95\%\ CI_{\bar{x}_m}$ 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 ($\bar{x}_m$ = 0.42) and for studies with indirect HRI ($\bar{x}_m$ = -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 ($\bar{x}_m$ = 0.65) than when taking part in indirect HRI ($\bar{x}_m$ = 0.03) or no HRI ($\bar{x}_m$ = 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 ($\bar{x}_m$ = 0.18) than participants in studies where the contact with the social robots was indirect ($\bar{x}_m$ = -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.

2.3.8.2 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).

2.3.8.3 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 2.1). 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.

2.3.8.4 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 2.1).

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 ($\overline{x}_m$ = 0.57) were more positive than were participants' attitudes from studies conducted in Germany ($\overline{x}_m$ = 0.22), Japan ($\overline{x}_m$ = 0.21), and the USA ($\overline{x}_m$ = 0.05).

Participants from studies conducted in France ($\overline{x}_m$ = 0.35) reported more positive cognitive beliefs about social robots than did participants who took part in studies conducted in Japan ($\overline{x}_m$ = 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 ($\overline{x}_m$ = 0.23) tended to be more positive than those of studies conducted in the USA ($\overline{x}_m$ = -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 $\overline{x}_m$.

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.

2.3.8.5 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 ($\overline{x}_s$) 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, $F(1, 20) = 1.35$, $p = .259$.

2.3.8.6 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, $F(1, 13) = 5.89$, $p = .031$.

2.3.8.7 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

($\bar{x}_s$) 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, $F(1, 28) = 0.001$, $p = .986$.

2.3.8.8 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 Appendix B). 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 ($\bar{x}_s$) 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, $F(1, 31) = 0.37$, $p = .549$.

2.4 Discussion

This review quantified and synthesised evidence on people's beliefs about social robots. Although reviews have been conducted in this area (Broadbent et al., 2009; Chen & Chan, 2011; Hancock et al., 2011; Savela et al., 2018), 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.

2.4.1 What are people's attitudes toward 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 (Ostrom 1969; Breckler, 1984). 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., 2006; Backonja et al., 2018; Rantanen et al., 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 1.1), 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 (de Graaf & Allouch, 2013; Zlotowski et al., 2015; Wullenkord et al., 2016) and sometimes with the purpose of predicting behaviour (Nomura et al., 2008; Park & Del Pobil, 2012; de Graaf & Allouch, 2013; Zlotowski et al., 2015; Wullenkord et al., 2016; Spence et al., 2018), attitude-behaviour models from social psychology should be used more consistently to inform HRI research (Pettigrew, 1998; Pratto et al., 2006; Hewstone & Swart, 2011).

2.4.2 To what extent do people accept, trust, and feel anxious toward 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 (Ostrom, 1969; Breckler, 1984).

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 2.4.3.

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 (Torta et al., 2013; Rosenthal-Von Der Pütten & Krämer, 2014; Hosseini et al., 2017), was used in 45% of the studies measuring anxiety and may have contributed to the overall neutral to positive valence for anxiety and trust.

2.4.3 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., 2007; Wullenkord et al., 2016; Bazzano & Lamberti, 2018). 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.

2.4.4 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 (Ebel, 1986; Broadbent et al., 2012). 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.

2.4.5 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 (Leite et al., 2013; Syrdal et al., 2014; Kachouie 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.

2.5 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). Such complexity may therefore warrant the use of approaches associated with the research of attitudes toward social out-groups such as Lord and Lepper’s (1993) Attitude Representation Theory which was the focus of the proceeding chapter.

Chapter 3: The Social Representation of Robots

Chapter Summary

Chapter 3 presents three studies that aimed to investigate how people, both on individual and group level, represent (or think about) robots with reference to two frameworks: Attitude Representation Theory (Lord and Lepper, 1993) and Social Representations (Abric, 1996). Study 2 was a small exploratory survey that offered initial evidence for the diverse, , ways in which people think about robots. Study 3 built on the results presented in Study 2 via an online survey in which participants were asked to indicate what comes to mind when they think about robots. The data was then used to construct a semantic network which reflected the representation of robots at a group (i.e., social) level using a UK-based sample. The social representation of robots was characterised by five modules of meaning (e.g., robots in their role as machines) that were further investigated in Study 4 which demonstrated that people's representation of robots at an individual level have a role in shaping people's attitudes toward robots. Furthermore, all three studies in Chapter 3 provided evidence for the salience of fictional depictions of robots as common exemplars of the robot category. Study 4 also demonstrated that such fictional depictions likely play a role in the formation of people's individual representation of robots.

3.1 Introduction

As discussed in Chapter 1, it is important to understand how people think and feel about robots. One challenge to answering this question is the wide variety of robotic systems and their lack of availability to the general public, which means that people may represent robots differently, with consequent effects on their attitudes. This observation could help to explain variability in attitudes toward robots in recent research (as reviewed in Chapter 2). It is therefore crucial to understand what comes to mind when people are asked about robots – or, in scientific terms, how people represent robots.

3.1.1 Attitude Representation Theory

As discussed in Section 1.3, Lord and Lepper's (1999) Attitude Representation Theory suggests that the way that people represent an attitude object likely influences their attitudes toward that object. Chapter 3 deals exclusively with the first postulate of this theory – *representation*. According to Lord and Lepper, an individual's evaluation of an attitude-relevant object (e.g., a specific robot) on the mental representation that the individual already has of the object and the broader category to which it belongs. This is particularly relevant in cases where individuals are asked to make broad evaluations of categories (e.g., robots in general), because it suggests that individuals may rely on pre-existing (potentially subjective) representations that they have of said category. Lord and Lepper suggest that people's representations are based on their personal experiences, thought processes, and subjective view-points. This has interesting implications in regards to robotics due to the previously noted rarity with which robots are encountered in real-life. Although there is substantial variety in the appearance, application, and function of real-life robots, it is likely that the general population has more experience with the relatively more homogeneous robots portrayed in fiction (Kriz et al., 2010). This may mean that, in general, individuals' representation of robots may not only be predominantly based on fictional robot, but there may also be a level of consensus between individuals' representations consistent with the characteristics of fictional depictions of robots.

3.1.2 Social representations

While Lord and Lepper predominantly discuss representations as specific to *each individual*, Abric (1996) defines social representations as the collective construction and transformation over time of representations by a given group. This distinction marks a recurrent difference between theories of attitudes and theories of social representations (De Rosa, 1993; Howarth, 2006) and has implications for the relationship between representations and attitudes that are not covered by the Attitude Representation Theory. Namely, the role of the social environment and the way individuals influence and are in turn influenced by it. This is particularly relevant to robots and the proposed impact of fiction on how people view and think about robotics since fiction can be used to shape meaning, portray collectively held beliefs and values, and influence social 'truths' about groups and objects (Lamarque & Olsen, 1996). Abric proposes that social representations consist of two components with distinct properties and functions – termed the central and peripheral systems. As the name suggests, the central system represents the 'core' that defines all aspects of the social category being represented. This core is not only resistant to change over time but also homogeneous across individuals as it is largely determined by information available to a specific social group (e.g., the general public in a given country). Such information encompasses long-term historical, cultural, and social conditions to which a given group is exposed. Due to this, the central system is not representative of, or responsive to, the immediate context within which the representation is relevant. In relation to robots, this particular property of the central system may imply that upon encountering a novel robot in real-life (e.g., MiRo, a zoomorphic robot), the core of the representation of robots an individual has: (a) is unlikely to change; and (b) may not be consistent with the properties of the novel robot (e.g., individual's representation may be defined by the idea that robots are humanoid). In contrast to the central system, the periphery of a social representation is flexible over time and can vary between people depending on their individual and subjective experiences. The periphery of a representation is responsive to novel information and contexts within which the representation is relevant (e.g., encountering a previously unseen robot), thus 'protecting' the core of the social representation by allowing for the integration of contradictory to the core information.

3.1.3 Measuring representations

Neither of the frameworks for understanding how people represent attitude objects (i.e., Abric, 1993; Lord & Lepper, 1999) suggest how to measure said representations. However, free associations are commonly used (e.g., Wagner et al., 1996; Smith & Joffe, 2013; Dany et al., 2015; Idoiaga et al., 2017). Free association is a fairly simple method where participants are asked to say or write down the first words that spontaneously come to mind after being presented with a concept cue (e.g., a word, picture, story, etc.). The resulting associations do not equate to the social representation as a whole but instead are considered as a method to describe the representation and identify the semantic elements which are accessible to the individual and their group and how they connect to each other (Lahlou & Abric, 2011). There are a number of ways in which the resulting associations can be analysed, typically with reference to the frequency with which the associations occur and their rank relative to other associations (Dany et al., 2015). Rank is considered a particularly important property as it assumed that the order in which associations occur is reflective of their importance in individual representations of the socially-relevant category (Dany et al., 2015). In other words, it is assumed that the first associations made by participants are the strongest elements of their individual representations of the given cue concept.

3.1.4 How do people represent robots?

Piçarra et al. (2016) asked 212 Portuguese adults to write the ideas (names, adjectives, etc.) that pop up into their mind when they hear the word robot. Following a methodology based on the frequency and rank of associations, Piçarra et al. divided the associations into four-quadrants, with each quadrant representing a part of the core and peripheral structures of the social representation. They found that "machine" was at the core of the representation as it was the association that was made most frequently. The periphery of the representation contained associations pertaining to the characteristics of robots such as "metal", "artificial intelligence", and "electronics"; the consequences of robotics such as "unemployment", "facilitates", and "innovation"; and references to specific robot groups such as "domestic robots" and "industrial robots". Piçarra et al. concluded that, as expected, the peripheral structure of the social representation of robots contained contradictory ideas (e.g., "unemployment" vs. "help") that reflect the variation in appearance and function of robots, but that the representation is ultimately defined by the fact that robots are "machines" (i.e., artificial and without emotions). Piçarra et al. further investigated the structure of the social representation by applying graph theory in order to understand the relationship between different elements of the representation (i.e., associations). For example, they found that "technology" co-occurs frequently with "help" just as "facilitates" is closely connected to "replaces men". Surprisingly, "machine" which represented the core of the representation was only weakly connected to only two other associations, "computer" and "technology".

While Piçarra et al.'s method offered a visual representation of the structure of the core and periphery of the concept of "robots", more advanced graphical methods can be used to analyse and visualise word associations in the context of social representations, namely the construction of semantic networks using complex algorithms (Palla et al., 2005; Keczer et al., 2016). Semantic networks have a number of advantages over the method used by Piçarra et al. First, analyses based on word frequency and rank order (as used by Piçarra et al.) require that qualitatively rich data is reduced to a single word or concept, which necessitates some level of interpretation from the researchers. In contrast, the construction of semantic networks using complex algorithms as described by Keczer et al. instructs participants to provide five words, thus removing the need for interpretation of the data. Second, constructing a semantic network following Keczer et al. procedure: (a) relies on normalised values rather than only frequency to ascertain the importance of different associations; (b) uses algorithms, which are a more sophisticated method of evaluating and describing the relationship between elements of the representations; (c) allows for statistical examination of the extent to which elements in the network are randomly organised (i.e., do not have the structure expected of a semantically described social representation) or not (i.e., there is a structure that would be expected if a central and peripheral system was present); and (d) allows all connections between all elements of the social representation to be visualized, thus providing a description of the entire representation and its structure. The studies presented in this chapter therefore aimed to expand on the work of Piçarra et al. by taking a more sophisticated approach to data analysis.

3.1.5 Research questions

In addition to providing a sophisticated assessment of how people represent robots, the studies presented in this chapter provide two important advancements over previous research on how people represent robots. First, the studies presented in this chapter explore the salience of fictional and non-fictional representations of robots. It is important to understand what robots people are most likely to think of as there is evidence that the way robots are represented in fictional media may affect people's attitudes and behaviour toward real robots (Kriz et al., 2010; Lorenčík et al., 2013; Mubin et al., 2015). This could be problematic as portrayals of robots in fiction rarely reflect the reality of current technology (Kriz et al., 2010) which may lead to some unrealistic expectations about robots. Whilst many studies have looked at people's attitudes toward robots in different contexts (Bartneck et al., 2007; Broadbent et al., 2007; Nomura et al., 2006) it is still unclear what role fictional representations of robots play in the formation and measurement of those attitudes. Second, the studies in this chapter aimed to establish a link between social representations and attitudes toward robots which, to my knowledge, has not been directly explored so far. If indeed people represent robots in varied ways, it would be expected that such representations would influence people's attitudes in accordance with Lord and Lepper's (1999) Attitude Representation Theory.

3.2 Study 2 - What Comes to Mind When People Are Asked Questions About Robots?

Study 2 aimed to provide some initial understanding of the way the concept of robots is represented by the general public, using relatively simple and quick to implement methodology that was based on free word associations in order to generate broad categories of meaning. Additionally, as previous work on attitudes toward robots identifies fictional representations of robots as a possible factor affecting people's expectations of robots' capabilities and thus their interaction with robotic devices (Broadbent et al., 2009; Kriz et al., 2010; Nomura et al., 2008), Study 2 also investigated the salience of fictional and non-fictional robots. It was expected that participants would name significantly more fictional robots than non-fictional robots.

3.2.1 Method

3.2.1.1 Participants

A short survey with 33 members of the general public was conducted at the Winter Gardens (an urban glasshouse open to the public) in Sheffield during working hours in March 2017.[4] Due to practical and ethical considerations, no personal information such as age and gender was collected. The majority of participants indicated that they did not work in an area related to robotics (N = 29, 87.9%) and that they have never visited Sheffield Robotics (N = 31, 93.9%), a prominent research centre in Sheffield.

3.2.1.2 Materials

Participants were asked four open-ended questions in the following order:

1. What comes to mind when you hear the word robot?
2. Can you list the first three robots that come to mind?
3. Do you work in an area related to robotics?

[4] A priori power analysis was not conducted for this study due to its explorative nature and sampling strategy.

4. Have you ever visited Sheffield Robotics?

These questions were selected as they were linguistically simple enough for people of various ages and backgrounds to understand. It was also thought that the first two questions were likely to induce relatively spontaneous responses in line with the free word association methods used in previous research. Participants' responses were recorded in a notebook. A printed out list of the questions was also kept and shown to participants who had difficulty hearing or understanding the questions. A box of chocolates was used as an incentive for participants to take part.

3.2.1.3 Procedure

Participants were approached as they were walking through the Winter Gardens and asked if they wished to take part in short survey. It was explicitly stated that no personal information would be collected and that the survey would take between two to three minutes. Individuals who looked under 18 years of age were not approached. Participants in groups or pairs were also avoided as it was difficult to know whether the presence of others may affect responses to the questions. Upon completing the survey, participants were offered a chocolate of their choice. Finally, participants were thanked for taking part, offered a debrief form that explained the purpose of the research, and asked if they had any questions about the survey (see Appendix D for debrief procedure).

3.2.1.4 Coding and analysis of word associations

For the first question, coding categories (i.e., nodes) were established prior to analysing the data but were ultimately not used as they were found to be unsuitable for coding participants' responses due to the variability (11 sub-categories) and relative length (M = 3.9 words per person) of the responses. Therefore, an approach based on manifest content analysis[5] was taken (Hsieh & Shannon, 2005). In other words, the analysis was largely descriptive, relying on the frequency of associations made by the participants, and involved little to no thematic coding or interpretation. This was deemed

[5] An essentially quantitative method. This is generally one of the first-steps toward a summative approach such as a latent content analysis (i.e., a more qualitative approach where there is a process of interpretation at the end of the analysis).

appropriate considering the spontaneous and context-free nature of participants' responses.

Before coding could take place, there was a period of familiarisation with the transcribed data which involved reading through participants' responses multiple times and noting any emerging categories. The data was then visually explored by generating a word cloud based on frequency using NVivo 11. The first step of the coding process involved going through each individual response (e.g., "Robot. Something mechanical, electrical.") and creating a new coding item (i.e., node) for each word or sentence that could be considered an individual category (e.g., "mechanical"). Words or sentences that were variants or synonyms of already established categories were coded using the existing nodes (e.g., "metal" and "metallic"). NVivo's Text Search function was also used to ensure accurate coding by conducting multiple word searches for stemmed and synonymous words. The final step of the coding process involved merging some of the nodes into single global categories where the separate nodes were essentially sub-categories. This was the most interpretive part of the coding process, although it was still predominantly guided by grouping together semantically similar words and sentences. It should be noted that semantic similarity here was defined by the researcher rather than by following one of the existing measures of semantic similarity.

3.2.1.5 Coding and analysis of fictional and non-fictional robots

As the second question aimed to establish whether people are more likely to think of fictional than non-fictional robots, a different approach to the one used for the word associations was used to code and analyse the data. Specifically, each entity mentioned by participants was coded as being either fictional, non-fictional, or other. The coding categories were developed prior to analysing the data and these have been described in Table 3.1. The response obtained from each participant was assigned a numerical value (0 – 4) for each of the three coding categories. For example, if a participant only said "Terminator", their response would have been given a value of 1 for the Fictional robot category and 0 for the other two categories. This coding approach is based on the assumption that an entity cannot be a member of more than one category. Finally, participants' responses were collated and the number of times each fictional and non-fictional robot was mentioned was counted and statistically compared.

Table 3.1

Coding Categories and Number of Items (n) Coded for Each Category

Category	Description	Examples
Fictional robot (36)	An entity that is the product of individual or collective imagination and does not physically exist in what is generally believed to be reality. The 'existence' of the entity is almost entirely constrained to an imaginary world depicting people, events, and places that are not factual.	Robocop WALL-E
Non-fictional robot (26)	An entity that is the product of individual or collective invention and physically exists in what is generally believed to be reality. The existence of the entity is almost entirely constrained to a world in which the people, events, and places are factual.	Asimo Robot Wars Drone
Other (10)	Any listed entity that cannot be coded under the 'fictional' or 'non-fictional' categories as described above.	Robotic man (meaning is unclear)

3.2.2 Results

Participants' responses were transcribed and coded using a mixture of NVivo 11 and Microsoft Excel 2016. Questions were coded using two distinct methods described in the corresponding sections below.

3.2.2.1 What does the general public associate with the word 'robot'?

The responses of 32 participants were coded as one participant was unable to provide an answer to the first question. A total of 39 nodes were coded with some participants (N = 4) providing a longer response containing multiple items. A total of 6 main categories and 8 sub-categories were established (see Table 3.2). The majority of associations (35.9%) reflected the artificial or non-organic features of robots, such as "metallic", "mechanical", and "artificial" (see Table 3.2). Furthermore, Artificiality, alongside Technology, are the two least variable categories relative to the number of nodes. In this case, variability is defined as the number of unique nodes (i.e., non-stemmed words and sentences) in each category divided by the total number of nodes. The closer this number is to zero, the less variability there is in the global category. For *Artificiality*, the variability number is 0.36 and for *Technology* the variability is 0.17. Compared to the next smallest variability, 0.67, for the *Technological Advancement* category.

3.2.2.2 Salience of fictional and non-fictional representations of robots

Following the initial coding, the total number of entities mentioned (n = 72) by each participant was recorded as was the total for each of the three categories. On average, participants named less than three entities (M = 2.18, SD = 1.11), with the majority of robots being fictional (n = 36, 50%), some non-fictional (n = 26, 36%), and only a few that could not be categorised as fictional or non-fictional (n = 10, 14%). There was no significant difference between the number of fictional and non-fiction robots mentioned by participants, $\chi^2(1) = 1.61$, p = .204. Further exploration of the data (see Figure 3.1) revealed that the most mentioned robot groups were robots from Star Wars (i.e., fictional robots; 12.5% of mentions) and industrial robots (i.e., non-fictional robots; 9.7% of mentions). A full list of all the entities mentioned by participants can be found in Appendix E.

Table 3.2

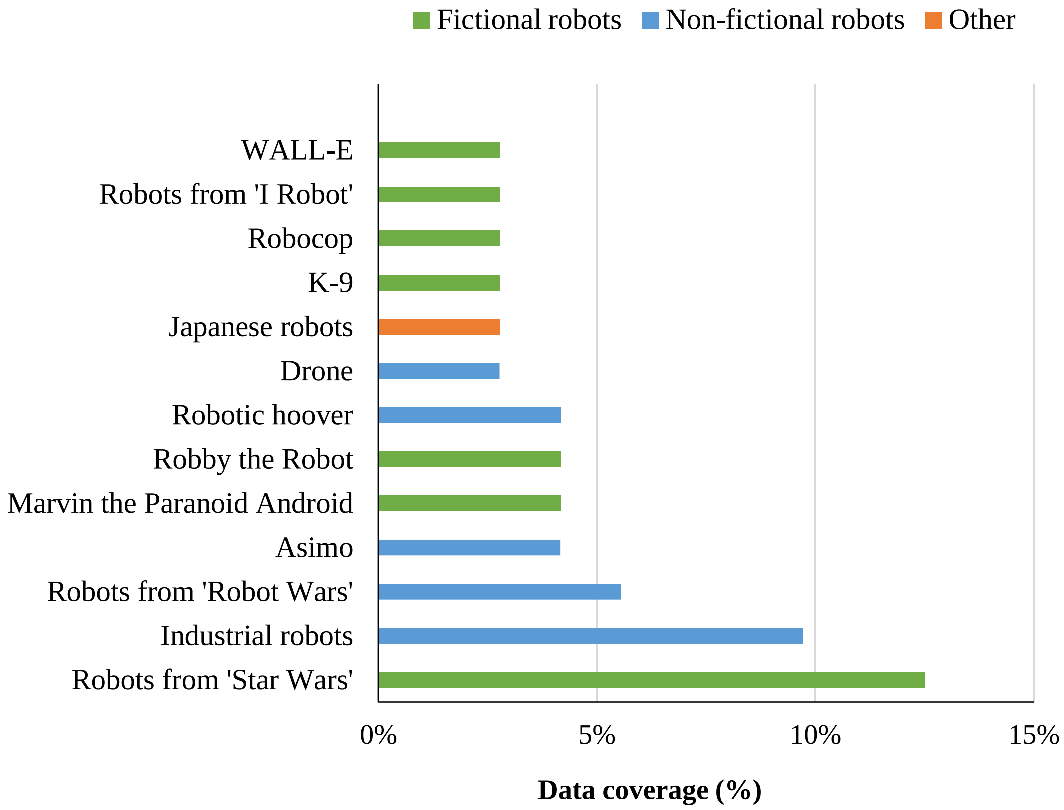
Coding Scheme for the Question "What Comes to Mind When You Hear the Word Robot?"

Category (*n*)	Sub-category (*n*)	Examples
Artificiality (14)	Mechanical (6)	"machine", "electrical"
	Metallic (4)	"metal", "metallic"
	Artificial intelligence (2)	"AI", "artificial intelligence"
	Artificial (2)	"artificial"
Technology (6)	-	"technology"
Media and Entertainment (6)	Television/Visual Media (4)	"Robot Wars"
	Other (2)	"the (robot] dance", "sci-fi"
Technological Advancement (6)	Technological progress (3)	"technology progressing", "automation"
	Future (3)	"the future", "futuristic"
Utility (5)	-	"carpet cleaner", "big machines used in factories"
Other (2)	-	"Universal Robots", "taking over the world"

Note. *n* denotes the number of items coded for each category.

Figure 3.1

Entities Mentioned by Participants that Occur in Over 2% of the Data



3.2.3 Discussion

In answer of the question "What does the general public associate with the word robot?", the results from the present study suggest that there is some thematic consistency in people's spontaneous responses. The most frequently made associations were the ones referring to the *artificial*, or otherwise non-biological, nature of robots. Predominantly neutral and descriptive words such as "mechanical", "metallic", and "artificial" were reported by participants and only one person gave a response which could be considered overtly negative ("[robots] taking over the world."). This is not necessarily surprising as both fictional and real robots are often, and at least in part, made from metal and can be defined as mechanical.

Other common associations made by participants were in reference to *technology* in general and only one participant referred to *automation* directly. This was somewhat surprising given the media coverage relating to robotics and automation at the time at which this study was conducted. In fact, more abstract associations (e.g., "the future") were fairly uncommon and participants mostly responded with specific examples (e.g., "big machines used in factories") or descriptive words (e.g., "artificial"). One possible explanation is that people's spontaneous responses are based on specific representations (or exemplars) from which they extrapolate information and that the visual attributes (e.g., being made from metal) of the robot are more readily brought to mind. Future studies may wish to consider whether participants' spontaneous associations correlate with the fictional and non-fictional representations of robots they can recall.

Given that the majority of the general population are unlikely to come into contact with robots on a regular basis, it was expected that participants would mention fictional robots more frequently than non-fictional ones as fictional exemplars of the robot category are arguably more common. This was the case but only marginally so, with 10 more mentions of fictional robots compared to non-fictional ones. This relatively small difference between the two categories could be for a variety of reasons including, but not limited to, the constraints provided by the question (i.e., only three robots), the quick pace of the survey (i.e., participants were not pressed to provide a minimum of three robot examples), and the relatively small sample size. It is also possible that experiences (even if rare) involving real robots are considerably more memorable, and thus more easily brought to mind, than experiences involving fictional robots. Further exploration of the data revealed that the most mentioned robot groups were industrial robots and robots from Star Wars. Whilst participants generally specified the robot characters from Star Wars, none were able to give any specific examples of industrial robots beyond specifying “car-making robots”. This pattern was observed for all of the fictional and non-fictional robots mentioned.

The findings from this study suggested that fictional representations of robots are frequently brought to mind when people are asked to spontaneously name robots and that regardless of fictionality, some exemplars of the robot category may be more predominant in the general population. While this study does not inform the relationship (if any) between people’s representations of robots and their attitudes, it does provide some incentive to consider the effect that fictional representations of robots may have on people’s attitudes and acceptance of robotics in general and suggests that Attitude Representation Theory may be a useful a way to explain the variability in attitudes toward robots (Enz et al., 2011; European Commission, 2012; Pino et al., 2015; Takayama et al., 2008).

However, Study 2 had a number of limitations that limit its usefulness in describing the social representation of robots in detail. First, the sample size was small and the participant recruitment strategy - combined with the fact that no demographic information was obtained - puts into question whether the sample was representative of the general public in the UK. Second, the method used to analyse the word associations was arguably sufficient to provide some initial insight into the social representation of robots but is ultimately too simplistic to provide a more comprehensive understanding of the representation’s structure as described by Abric (1993).

3.3 Study 3 - A Semantic Network of the Social Representation of Robots

The results of Study 2 suggested that there might be some thematic consistency in the associations people make with the word "robot" that is indicative of an underlying central and peripheral structure of the social representation of robots. Results from Study 2 also demonstrated that fictional representations of robots (e.g., C3P0 from Star Wars) are often brought to mind and that some representations of robots may be more predominant in the general population. Having said that, the survey in question had a number of limitations that have been addressed in Study 3: a larger sample was recruited, demographic information was collected, a more rigorous approach to analysing the data was taken, and participants' attitudes were measured using the Negative Attitudes toward Robots Scale (NARS; Nomura et al., 2004) and a human-robot version of the Implicit Association Test (IAT; Greenwald et al., 1998) developed by MacDorman et al. (2009). The aims of Study 3 were to: (a) construct a semantic network of the social representation of robots; (b) further explore the salience of fictional and non-fictional representations of robots; and (c) investigate whether there is a relationship between the number of fictional and non-fictional robots that participants mention and their attitudes.

3.3.1 Method

3.3.1.1 Participants

A sample of 106 fluent English speakers completed the online study via Qualtrics. The majority of participants (*N* = 65) were recruited via a mailing list of staff volunteers at the University of Sheffield and through a Facebook post advertising the study which was shared by members of the researcher's personal social network.[6] As a compensation for their time, participants were entered into a draw for a £25 Amazon voucher which was randomly awarded after the end of the data collection period. First year psychology undergraduates (*N* = 41) were also recruited via the Department of Psychology's Online

[6] A traditional priori power analysis was not conducted for this study but a sample of *N* = 100 was determined priori as sufficient for constructing a semantic network while still being a feasible sample to recruit with limited resources.

Research Participation System (ORPS). These students received 1 course credit as compensation for their time. Table 3.3 provides details of the sample. It should be noted that 5 participants only completed part of the study and as such were not included in all of the analyses described in the Results section (this is explicitly stated in the relevant sections). Additionally, one participant was excluded as their answers to the first two tasks put doubt as to whether they have completed the study correctly.

Table 3.3

Participants' Demographic Characteristics

Characteristic	
Age	
M	30.40
SD	12.19
Range	18 – 64
Gender	
Female (%)	73 (68.87)
Male (%)	27 (25.47)
Non-binary (%)	1 (0.94)
Not specified (%)	5 (4.72)
Nationality	
UK	75
Not specified	9
China	2
Romania	2
Turkey	2
USA	2
Belgium	1
Bulgaria	1
Chile	1
El Salvador	1
Estonia	1
Iran	1
Ireland	1
Italy	1
Jordan	1
Malaysia	1
New Zealand	1
Poland	1
Portugal	1
UK/USA	1
Studied or worked in an area related to robotics	
No (%)	91 (85.85)
Yes (%)	10 (9.43)
Not specified (%)	5 (4.72)

Note. Four participants did not complete the entire study and as such demographic data is missing for them. One participant was excluded and their demographic data was removed. The age demographics are thus based on 101 participants instead of the full sample.

3.3.1.2 Procedure and Materials

Listing and identifying robots. After giving informed consent, participants were shown a message warning them that the task that they were to complete first was timed and that they should only proceed if they could dedicate 1 minute and 30 seconds to said task. After confirming that they had read this message, participants were asked to write down the first five robots that came to mind. Participants were also instructed to state where they have seen the robot or to briefly describe it if they could not remember its name. Participants were automatically moved onto the next part of the task after 1 minute and 30 seconds had elapsed, regardless of whether they managed to list five robots or not. Following this, participants were asked to indicate whether the robots that they wrote down, if any, were fictional (i.e., cannot be found in real life), non-fictional (i.e., can be found in real life), or were neither/both. This part of the task was not timed and participants' answers from the first part of the task were automatically embedded in the multi-choice questions.

Association task. Participants received the same message as in the previous task, indicating that they would only have 1 minutes and 30 seconds to complete the next task. Following the methodology described in Keczer et al. (2016), participants were asked to write down the first five words that came to mind when they thought about "robots". They were asked to do so in separate text boxes and were moved onto the next part of the task automatically, regardless of whether they were able to list five associations or not. Following this, participants were asked to indicate the extent to which the words or phrases they wrote down had a positive, neutral, or negative meaning to them in relation to the word "robot". This part of the task was not timed and participants were told that there were no right or wrong answers.

Implicit Association Test (IAT). Participants were then asked to complete an IAT (Greenwald et al., 1998) – a computer task that has been developed to measure the strength of associations between different pairs of concepts. For the purposes of this study, MacDorman et al., (2009) robot-human IAT was adapted following the guidance provided by Greenwald et al. (2003) to assess participants' implicit attitudes toward robots, relative to humans. This particular version of the IAT consisted of seven testing blocks. For all blocks, the image and word stimuli appeared on screen until the participant responded. Category labels appeared at the top left and right corners of the screen. If participants classified a target stimulus incorrectly, a red cross appeared bellow the stimulus until the correct response was made. There was an interval of 400ms between each trial. The free tool *iatgen* (iatgen.wordpress.com) was used to create an IAT compatible with Qualtrics based on the procedure detailed in Greenwald et al. (2003). Participants were first asked to classify ten human and ten robot silhouettes as either *human* or *robot* by pressing the E (left) or I (right) keyboard keys. They then completed the second block of the task by classifying 16 words as either *pleasant* (good, joy, love, peace, wonderful, pleasure, friend, laughter, happy) or *unpleasant* (agony, terrible, horrible, nasty, evil, war, awful, failure) using the E (left) or I (right) keyboard keys. In the third and fourth blocks, participants categorised both the images and words in the same manner. For the fifth testing block, participants were asked to repeat the second block (i.e., categorising pleasant and unpleasant words) but with the position of the labels (pleasant vs. unpleasant) reversed. In the sixth and seventh blocks, participants were once again asked to categorise both the images and words but with the pairings reversed. Participants were instructed to be as quick and as accurate as possible. The starting position (left or right) and the combination of the words and images (pleasant-human vs. unpleasant robot and pleasant-robot vs unpleasant-human) was randomised across participants.

Participants' individual *d*-scores were generated automatically from the Qualtrics output via the free tool iatgen (iatgen.wordpress.com) and entered for analyses in SPSS. The minimum *d*-score that could be achieved was -2 and the maximum was +2. Positive scores indicate that participants were quicker to make more positive associations with the human silhouettes than they were with the robot silhouettes. The opposite is true for negative *d*-scores. A *d*-score of 0 would potentially indicate that participants were not biased toward either group.

Negative Attitudes toward Robots Scale (NARS). Following the IAT, participants were asked to complete a modified version of the Negative Attitudes toward Robots Scale (NARS) to measure their explicit attitudes toward robots (Nomura et al., 2004; Tsui et al., 2010). The NARS is comprised of three subscales (*Interaction with robots*, *Social influence of robots*, and *Emotion in interaction with robots*) with a total of 16 items (see Table 3.4). Each item is a statement that can be rated on a five-point scale from 1 (strongly disagree) to 5 (*strongly agree*). The order in which the items were presented was randomised. The reliability and validity of the NARS has been supported by multiple studies (Nomura et al., 2004; Nomura et al., 2006).

Participants' mean for each of the three subscales of the NARS was generated by summing up the score for each of the items in the *Interaction*, *Social influence*, and *Emotion* subscales as per the instructions in Nomura et al. (2004). The scores of the Emotion subscale were inverted prior to summation. For the Interaction subscale, the minimum score is 5 and the maximum score is 25 (after the removal of an item); for Social influence the minimum is 4 and the maximum is 20 (after the removal of an item); and for the Emotion subscale, the minimum is 5 and the maximum is 25. Larger values indicate more negative attitudes, while smaller values indicate more positive attitudes.

Table 3.4

Items and Cronbach's α Values for Each Subscale of the Negative Attitudes Toward Robots Scale (NARS)

Subscale	Items	Cronbach's α
NARS-S1: Interaction with robots	I would feel uneasy if I was given a job where I had to use robots.	.80*
	The word "robot" means nothing to me. [a]	
	I would hate the idea that robots or artificial intelligences were making judgements about things.	
	I would feel nervous operating a robot in front of other people.	
	I would feel nervous just standing in front of a robot.	
	I would feel paranoid talking with a robot.	
NARS-S2: Social influence of robots	I would feel uneasy if robots really had emotions.	.74*
	Something bad might happen if robots developed into living beings. [a]	
	I feel that if I depend on robots too much, then something bad might happen.	
	I am concerned that robots would be a bad influence on children.	
	I feel that, in the future, society will be dominated by robots.	
NARS-S3: Emotion in interaction with robots (inverse)	I would feel relaxed talking with robots.	.82
	If robots had emotions, then I would be able to make friends with them.	
	I feel that I could make friends with robots.	
	I would feel comfortable being with robots.	
	I would feel comforted being with robots that have emotions. [b]	

[a] Denotes items that were removed in order to improve the overall reliability of the scale.

[b] Denotes items that have been adapted into a 'would' statement.

*Cronbach's α after item deletion.

Demographic questions. At the end of the study, participants were asked to provide their age, gender, and nationality (or state that they "Prefer not to say."). They were also asked if they have ever studied or worked in an area related to robotics and what their level of personal interest in robotics was on a scale from 1 (*Completely disinterested*) to 5 (*Extremely interested*). Finally, participants were asked if they wished to be entered into a prize draw (or were awarded 1 credit if Psychology undergraduates), debriefed, and allowed to leave additional comments (see Appendix D for debrief procedure).

3.3.1.3 Construction of the semantic network

Visualising social representations using semantic networks. Networks can be broadly defined as mathematically generated structures (i.e., graphs) that represent the relationships between pair-wise objects (e.g., neurons, cells, social groups, etc.). A semantic network is one that aims to model any part of language. As Lahlou and Abric (2011) point out, "language is a network where each term is defined by other terms" (pg. 5) which is relevant to the idea of describing social representations by examining the occurrence and relationship between word associations in a semantic network. Furthermore, semantic networks made of word associations have properties that can reveal something about the central and peripheral structures of social representations. Namely, the majority of objects (i.e., words) in large semantic networks occur relatively rarely and tend to have only a few links to other objects in the network with only a few hubs of greater connectivity (Steyvers & Tenenbaum, 2005). This structure parallels the heterogeneous nature of the peripheral system which is defined by individuals' personal experience rather than collective societal contexts which are responsible for the core of social representations.

Using the data obtained from the association task, an undirected weighted semantic network representing the concept of "robots" (termed the *cue concept*) was constructed by following the methodology used by Keczer et al. (2016). An undirected network is one in which the relationship (i.e., edges) between the objects (i.e., nodes) in the network, in this case individual associations with the word 'robots', is symmetrical and unidirectional. A weighted network is one where the edges are assigned some value, typically a measure of the strength of the relationship between the nodes. The data from 105 participants was used to identify the nodes (i.e., individual associations with the cue concept) and edges (i.e., the connections between associations) which formed the basis of the semantic network. The following procedure was undertaken in the order of steps that is presented below. The resulting network was visualised using the ForceAtlas 2 layout in Gephi 0.9.2.

Assessing and cleaning of the data set. Prior to identifying the nodes, the data was checked for any repetition of associations at the participant level – e.g., a participant's first and third associations were both "metal". Repetitions were excluded from the semantic network if found. Where words were misspelled but still intelligible (e.g., "metalic" instead of "metallic"), the data was retained and later coded under the correct spelling. There were only 2 associations made by two different participants that were completely unintelligible (namely, "montotoned" and "monot") and as such were removed from the data set without excluding the participants' other responses. No further deletions were made.

Identifying and coding the nodes. NVivo was used to identify, code, and count nodes. Each node represents an association with the cue concept, in this case *robots,* that has been made more than once by two or more different participants. As such, all associations that occurred only once were excluded from the semantic network in line with the methodology used by Keczer et al. (2016). Synonymous words were merged into a single node and the node was coded under the word that occurred most often (e.g., the associations "quick" and "fast" were coded under the node: *Fast*). Synonyms were identified using NVivo's 'Text Search' function. Associations consisting of multiple words (e.g., "science-fiction" or "science fiction") and their popular abbreviations (e.g., "sci-fi") were coded under the same node, using the abbreviated version to ease the visualisation of the network later on.

Determining the edges and edge weights. An edge is a measure of the co-occurrence of two associations at the participant level. For example, if a single participant associated the cue concept of 'robots' with both "metal" and "artificial", there was considered to be an edge between this particular pair of associations. Edges were identified using NVivo and recorded in Excel 2016 along with the edge weights. The weight (*w*) of each edge is equal to the number of times that a pair of associations occurs together. For example, if five different participants associated the concept of 'robots' with both "metal" and "artificial", then the edge weight of this association pair would be five.

Calculating the degree, strength and normalised strength of nodes. For an undirected weighted network, the following equations were used to calculate the degree (k_i), strength (s_i), and normalised strength ($s_{i_{norm}}$) of each node.

The degree of a node *i* is defined as the total number of edges (*a*) connected to it in the network (*j*):

$$k_i = \sum_j a_{ij}$$

While the strength of a node *i* is the sum of the weights (*w*) of all edges connected to it:

$$s_i = \sum_j w_{ij}$$

Finally, the normalised strength of node *i* is its strength (s_i) divided by the average node strength of the network ($\bar{s}$):

$$s_{i_{norm}} = \frac{s_i}{\bar{s}}$$

Determining whether the network is scale-free. Broadly defined, a scale-free network is one where the distribution of edges from a given node (i.e., the node's degree) follows (i.e., can be plotted as) a power-law distribution (Crucitti et al., 2003). This should be true for some, but typically not all, nodes in a given network, meaning that a network can be scale-free to a greater or lesser extent. The properties of scale-free networks differ from random networks in a number of ways which guide how the network is constructed, analysed, and visualised, such as: (a) global hubs are present within the network (see next subsection); (b) new nodes added to the network are more likely to link to existing nodes that have a larger degree; (c) the network is less likely to collapse if nodes are taken out at random but more likely to collapse if one of the global hubs is taken out. Simply put, the structure of network would be notably changed such that there are a substantial number of nodes that are not connected to each other. Large (i.e., consisting of approximately a 100 nodes) semantic networks made of word associations tend to be scale-free (Steyvers & Tenenbaum, 2005) and therefore not random. This is important in the context of social representations as defined by Abric as we would expect to see distinct structures (i.e., modules) in the network that would signify the presence of a central and peripheral systems of the social representation.

In order to determine whether the social representation network of 'robots' was scale-free, the normalised node strengths ($s_{i_{norm}}$) of the network were fitted to a power-law distribution using the Matlab functions provided here: http://tuvalu.santafe.edu/~aaronc/powerlaws/. The scaling parameter (α) of the power-law function of the network, minimum value (x_{min}) of $s_{i_{norm}}$ for which the power-law holds true, and the Kolmogorov-Smirnov p-value for the fitted power-law model were calculated. Next, 2500 datasets with the same α and x_{min} that follow the power-law distribution were generated and the Kolmogorov-Smirnov p-value for each dataset was calculated. In order to test whether the fit of the data to the power-law distribution was plausible, the distance between the distribution of the empirical and generated datasets was compared (Clauset et al., 2009). The resulting p-value quantifies the plausibility of the fit and is defined as the fraction of the distances that are larger than the empirical dataset. It should be noted that if p is close to 1, then the difference between the empirical data and the generated datasets can be attributed to random factors and the power-law is a plausible fit for the empirical dataset. Following the methodology described by Keczer et al. and Clauset et al., the network was determined to be approximately scale-free (see Results for details).

Identifying the global hubs of the network. Global hubs are broadly defined as the nodes in a scale-free network that represent the most dominant associations with the cue concept, in this case, the word "robots". For such hubs, the number of edges connected to the node should exceed the average number of edges for all nodes in the network given a specific value (termed the "selection threshold"). In a scale-free network, it would be expected that the larger the number of nodes, the bigger the global hubs would be as new nodes are more likely to connect to the hubs than any other node in the network. The selection threshold for the global hubs was based on the normalised node strengths ($s_{i_{norm}}$). The $s_{i_{norm}}$ values of all nodes were ranked such that the largest values came first. Following this, the k_i of each node, its rank, and its corresponding $s_{i_{norm}}$ were then examined in order to decide the selection threshold for the global hubs. It was noted that nodes with ranks < 10 had considerably more edges than nodes with lower ranks and that the majority of nodes had $s_{i_{norm}} < 2$. As such, it was decided that any node with $s_{i_{norm}} \geq 2$ would be considered a global hub for the network. Although this decision may appear somewhat arbitrary, it is in line with other methods of selecting threshold values for global hubs (Li et al., 2005).

Determining the modularity of the network. A Louvain community algorithm with fine-tuning was applied to the weighted network in order to determine the partition of the modules and the network's modularity value (Q) via Matlab (the script can be found here: https://sites.google.com/site/bctnet/). The Louvain community method (Blondel et al., 2008) detects communities (i.e., modules or clusters) with data that have a high density of links (i.e., edges) between nodes within the modules but as few as possible intermodular links. The modularity value (Q) reflects the quality of the partitioning of the modules and values close to 1 signify a network with a very high density of intramodular edges and very low density of intermodular edges. As the Louvain algorithm yields slightly different values of Q for each iteration, Keczer et al.'s (2016) methodology was followed in order to find the highest possible modularity value from 10,000 independent runs. The partition of the modules with the highest value of Q were selected.

In order to determine how plausible this modularity was, the network's maximal modularity (Q) was compared to the maximal modularity (Q_{rand}) of 100 random networks. One hundred independent random undirected networks with preserved weight, degree and, strength distributions were generated (the script can be found here: https://sites.google.com/site/bctnet/). A one-sample t-test was then applied in order to determine whether Q was always greater than the Q_{rand} that would be expected by chance for a random network.

Identifying the modular hubs. Modular hubs are broadly defined as the nodes in each module of a network that represent the dominant association(s) for that particular module. For such hubs, the number of edges connected to the node within its own module should exceed the average number of edges for all the other nodes in the same module. Typically, the global hubs are also modular hubs. The selection threshold for the modular hubs was based on the normalised intramodular node strengths ($sm_{i_{norm}}$) and followed the same logic as the selection of the threshold for the global hubs. As such, any node within a module with $sm_{i_{norm}} \geq 2$ was considered as a hub. The following equations were used to calculate the intramodular strength (sm_i) and normalised intramodular strength ($sm_{i_{norm}}$) of each node.

The intramodular strength of a node (i) is the sum of the weights (w) of all edges connected to it within its own module (m):

$$sm_i = \sum_m w_{im}$$

The normalised intramodular strength of node (i) is its intramodular strength (sm_i) divided by the average intramodular strength of all the nodes in the module ($\overline{sm}$):

$$sm_{i_{norm}} = \frac{sm_i}{\overline{sm}}$$

3.3.2 Results

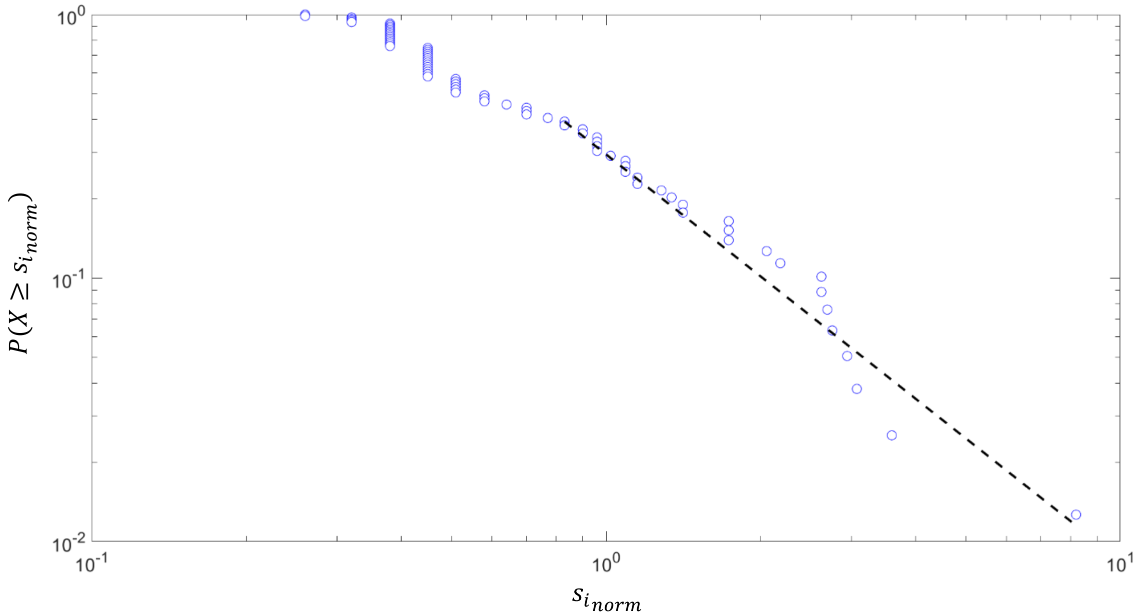
3.3.2.1 Semantic network of the social representation of robots

Properties of the network. A total of 517 associations with robots (positive associations = 175, negative associations = 77, and neutral associations = 265) made by 105 participants were considered for the semantic network (i.e., an average of 4.92 associations per participant). On average, participants made 1.66 (SD = 1.31) positive associations, 0.73 (SD = 0.85) negative associations, and 2.52 (SD = 1.31) neutral associations with robots. The number of associations that occurred more than once resulted in a network containing n = 79 nodes (full list of nodes can be found in Appendix F).

The scale free properties of the network were estimated as α = 2.54, x_{min} = 0.83, p = 0.184. The power-law distribution of $s_{i_{norm}}$ for the network's nodes is plotted in Figure 3.2. Comparison to 2500 generated datasets indicated that the power-law distribution was a plausible fit for the empirical data (p = 0.74). It should be noted that the α for all datasets were adjusted to account for the lower than desirable number of observations (i.e., number of $s_{i_{norm}}$ data points < 100). This was a 'finite-size' adjustment embedded in the Matlab function found here: http://tuvalu.santafe.edu/~aaronc/powerlaws/.

Figure 3.2

A Plot of the Empirical Distribution of the Normalised Node Strength ($s_{i_{norm}}$*) of the Network Along with the Fitted Power-Law Distribution (Dashed Line) on Log-Log Axes where the Y-axis is the Cumulative Density Function of the Normalised Node Strength.*



Global hubs of the network – the core of the social representation. Ten global hubs were identified and have been listed in Table 3.5. "Metal" is the hub with the largest normalised node strength that is approximately 3 *SD* away from the mean $s_{i_{norm}}$ of the 10 global hubs. This indicates that "metal" is a node with an outstanding number of connections to other nodes and as such can be considered as the dominant association that represents the concept of *robots* or, in other words, a key part of the core system of the social representation of robots. Half of the global hubs are associations that are descriptive in nature ("mechanical", "helpful", "emotionless", "artificial", and "useful"). The global hubs "AI", "machine", and "computer" can be interpreted as variations of the concept of 'robots' and do not appear to represent any abstract concepts. "Technology" is the only hub that appears to represent a broader category.

Table 3.5

List Of The Network's Global Hubs and Their Corresponding Degree (k_i), Strength (s_i), and Normalised Strength ($s_{i_{norm}}$)

Global Hub	k_i	s_i	$s_{i_{norm}}$
metal	58	128	8.26
AI	34	56	3.59
machine	30	48	3.10
mechanical	28	46	2.97
helpful	28	43	2.78
emotionless	29	42	2.71
artificial	28	41	2.62
technology	28	41	2.65
computer	23	34	2.19
useful	26	32	2.07

Modularity of the network – the organisation of the social representation. A one-sample *t*-test was conducted to determine whether the modularity value (Q) of the 'robots' network was significantly higher than the modularity (Q_{rand}) of 100 random networks. For all random networks, $Q > Q_{rand}$. The *t*-test was significant, p < .001, $M_{Q_{rand}}$ = 0.176 ($SD_{Q_{rand}}$ = 0.008), indicating that the semantic network representing *robots* is modular (i.e., has an underlying structure as would be expected for a model of a social representation). The modularity value of the network was $Q = 0.254$ and five separate modules were identified. The modules and modular hubs are listed and described in Table 3.6 (for a full list of all nodes in each module, see Appendix G). Figure 3.3 is a visualisation of the entire network. Figure 3.4 represents all the nodes and their average valence. A visual break-down of the modules can be found in Appendix H.

Module 1 (depicted in green in Figure 3.3) had the highest density of edges and the largest average intramodular strength (sm_i = 10) as well as the highest number of modular hubs ("metal", "AI", "computer"). With the exception of the hubs, the normalised intramodular strength ($sm_{i_{norm}}$) of the nodes in the module were fairly homogeneous (*SD* = 0.31 for the non-hub nodes; *SD* = 1.06 for the entire module). This finding further supports the modular hubs as the dominant associations for this module and suggests, in line with the properties of a scale-free network, that removal of these hubs would dramatically change the modularity of the network (and thus, the underlying structure of the representation of the concept of robots).

Module 2 (depicted in yellow in Figure 3.3) had the second highest density of edges and average intramodular strength (sm_i = 8.55). When the two modular hubs ("helpful" and "useful") were excluded, the normalised intramodular strength ($sm_{i_{norm}}$) of the nodes in the module became somewhat homogeneous (*SD* = 0.44 for the non-hub nodes; *SD* = 0.67 for the entire module). Similar to Module 1, this finding indicates that the hubs are indeed important for the structure of the network but their removal would be less impactful than the removal of the hubs of Module 1.

Module 3 (depicted in blue in Figure 3.3) had a similar density of edges and average intramodular strength (sm_i = 5.78) to Module 4. It did, however, contain two modular hubs ("emotionless" and "mechanical") without which the module's normalised intramodular strength becomes fairly homogeneous (*SD* = 0.43 for the non-hub nodes; *SD* = 0.77 for the entire module). These findings indicate that the associations that these modular hubs represent, are an important part of the representation of the concept of robots (but less so than the hubs of Module 1 and 2).

Module 4 (depicted in orange in Figure 3.3) had the second lowest density of edges and average intramodular strength (sm_i = 5.67). It has a single modular hub ("artificial") that, when removed, does not impact the homogeneity of the normalised intramodular strength ($sm_{i_{norm}}$) greatly (*SD* = 0.41 for the non-hub nodes; *SD* = 0.53 for the entire module).

Module 5 (depicted in pink in Figure 3.3) contained the least number of nodes, had the lowest edge density, and the lowest average intramodular strength (sm_i = 5). This module had no nodes that pass the $sm_{i_{norm}}$ threshold that would indicate the presence of modular hubs. In addition, the normalised intramodular strength of its nodes is fairly homogeneous (*SD* = 0.36). This finding may indicate that the associations in this module are peripheral to the core associations that represent the concept of robots.

Table 3.6

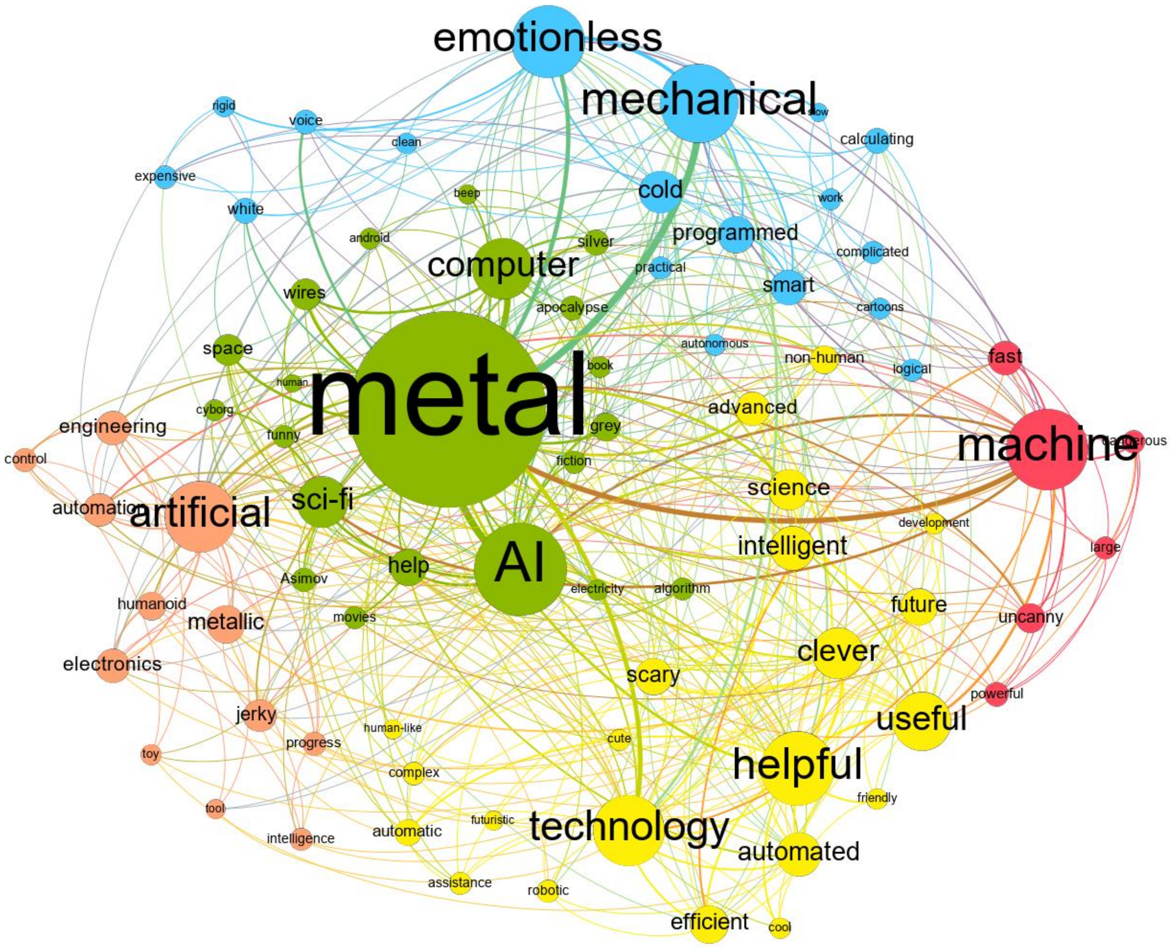
List of the Network's Modules and Modular Hubs, the Number Of Nodes in Each Module (n_i), and the Intramodular Strength (sm_i) and Normalised Intramodular Strength ($sm_{i_{norm}}$) of the Hubs

Module name (n_i)	sm_i	$sm_{i_{norm}}$
Module 1 (21)		
metal	46	4.60
AI	29	2.90
computer	23	2.30
Module 2 (22)		
helpful	22	2.57
useful	20	2.34
Module 3 (18)		
emotionless	18	3.11
mechanical	14	2.42
Module 4 (12)		
artificial	12	2.12
Module 5* (6)		
-	-	-

*All nodes of the module are below the threshold value and as such no modular hubs were identified.

Figure 3.3

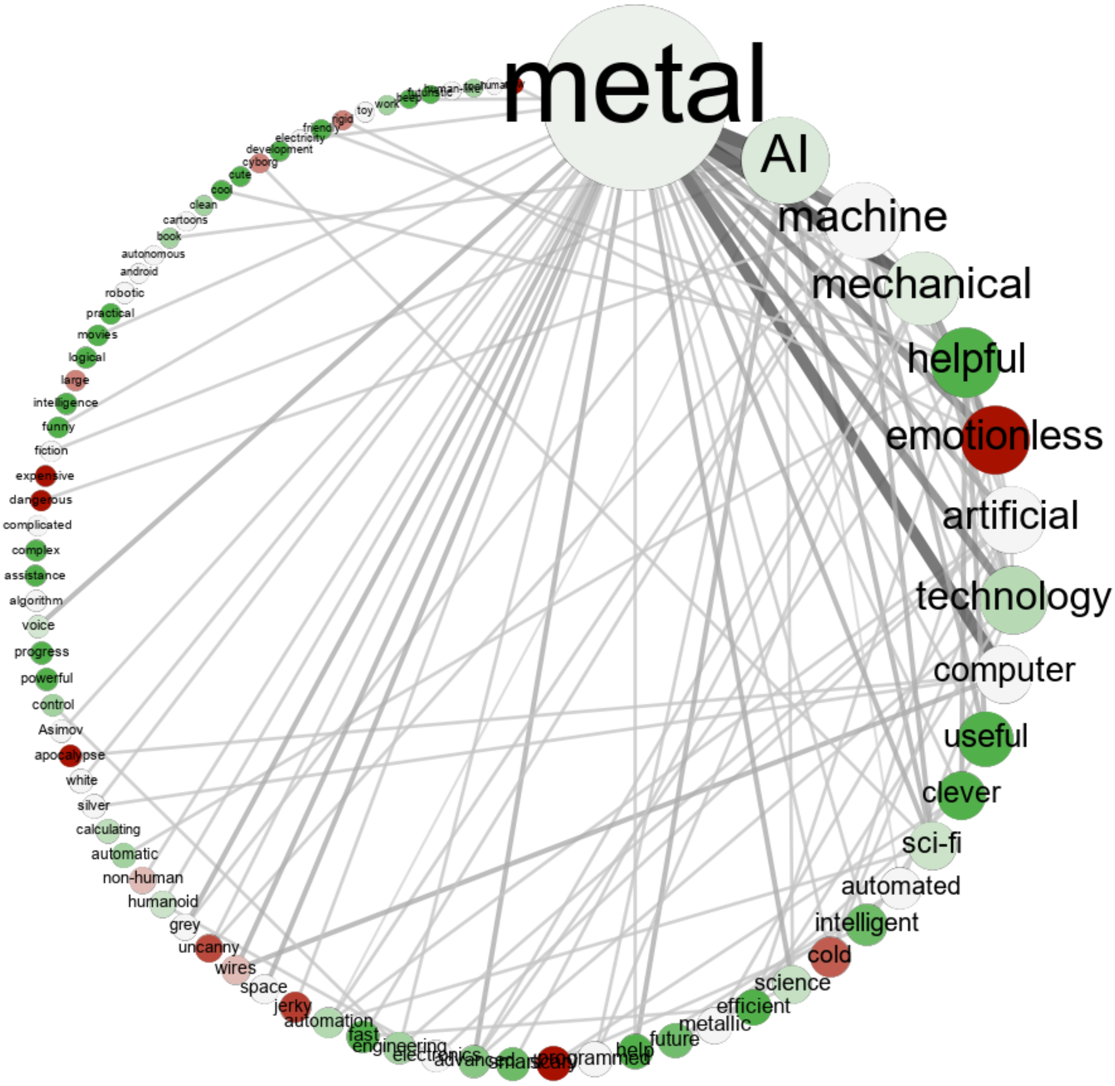
Visualisation of the Social Representation of Robots



Note. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of 1 have been removed for better visualisation. Module 1 (Green); Module 2 (Yellow); Module 3 (Blue); Module 4 (Orange); Module 5 (Pink).

Figure 3.4

Visualisation of all the Nodes / Associations in the Network and their Average Valence



Note. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

3.3.2.2 Relationship between the valence of associations and explicit and implicit attitudes

Five participants did not complete the NARS and as such were excluded from the analysis in this section. Therefore, the results reported below are based on the data of 100 participants rather than 105 participants. A regression was conducted to see whether the valence of the associations which participants made predicted each of the four dependant variables: attitudes toward interaction with robots (NARS-S1), attitudes toward the social influence of robots (NARS-S2), attitudes toward emotion when interacting with robots (NARS-S3); and implicit attitudes as measured via the IAT (*d*-scores). Statistical assumptions were checked for each regression to ensure that there was an independence of residuals, a linear relationship between the dependant and independent variables, homoscedasticity, approximately normally distributed residuals, and no outliers for all tests. Participants' ratings of the valence (i.e., positive, neutral, or negative) of each of the word associations they made was summed such that an overall valence between +5 and -5 was calculated for each participant. It should be noted that no adjustments were made to account for participants who did not provide five associations and as such the ratio of positive to negative to neutral associations was not taken into account. On average, the overall valence of participants' associations was slightly positive (M = 0.99, SD = 1.70, range = -3–5).

The overall valence of the associations significantly predicted participants' attitudes toward interaction with robots (see Table 3.7) and accounted for 11.1% of the variance in participants' NARS-S1 scores, $F(1, 98) = 12.26$, $p = .001$, $R^2 = 0.11$. The overall valence of the associations also significantly predicted participants' attitudes toward the social influence of robots (see Table 3.7) and accounted for 10.9% of the variance in participants' NARS-S2 scores, $F(1, 98) = 12.02$, $p = .001$, $R^2 = 0.11$. The same was also true for participants' attitudes toward emotion when interacting with robots (see Table 3.7) and the valence of the associations accounted for 12.5% of the variance in participants' NARS-S3 scores, $F(1, 98) = 13.95$, $p < .001$, $R^2 = 0.13$. However, the valence of the associations did not predict participants' implicit attitudes as measured via the IAT, $F(1, 98) = 0.02$, $p = .901$, $R^2 = 0.00$.

Table 3.7

Regression Coefficient (B), Standard Error of the Coefficient (SE_B), and Standardised Coefficient (β) of the Regression for NARS-S1, NARS-S2, NARS-S3.

	B	SE_B	*β*
NARS-S1			
Intercept	13.67	0.42	
Overall valence	-0.75	0.22	-.33
NARS-S2			
Intercept	13.43	0.34	
Overall valence	-0.60	0.17	-.33
NARS-S3			
Intercept	15.99	0.41	
Overall valence	-0.78	0.21	-.35

3.3.2.3 Salience of fictional and non-fictional representations of robots and relationship with explicit and implicit attitudes

The data from 104 participants was used to assess the salience of fictional and non-fictional representations of robots. On average, participants mentioned a total of 4.06 (*SD* = 1.19) robots with the majority of participants (*n* = 53, 51%), being able to name or describe five robots as requested. Any inadequate responses (e.g., "I don't know", "N/A", etc.) were excluded from the data set. The approach to coding the data was similar to the one described in the Results section of the first survey (see Section 3.2.2.2) however, in the present study, it was the participants that were responsible for classifying the fictional status of the robots rather than the researcher. Before conducting any analyses, the extent to which there was agreement between the participants and the researcher in the way that the robots were classified as *fictional*, *non-fictional*, or *other* was checked. Specifically, the researcher second-coded each response using the definitions from the first survey (see Table 3.1). For fictional robots there was 94.23% agreement between the participants and the researcher, 91.35% agreement for non-fictional robots, and 88.46% agreement for other types of robots. Since it is not possible to say whether the disagreements resulted from human error (e.g., participants selecting the wrong label by mistake) or due to participants' genuine belief in the fictional status of the robots, it was decided that the classification that the participants provided would be used for analysing the data as it was more likely to represent participants' belief about robots and their fictional status. It should be noted that, on average, participants had neutral or slightly positive attitudes toward robots (NARS-S1, M = 12.92, SD = 3.85; NARS-S2, M = 12.84, SD = 3.09; NARS-S3, M = 15.21, SD = 3.78; IAT d-score, M = 0.54, SD = 0.34).

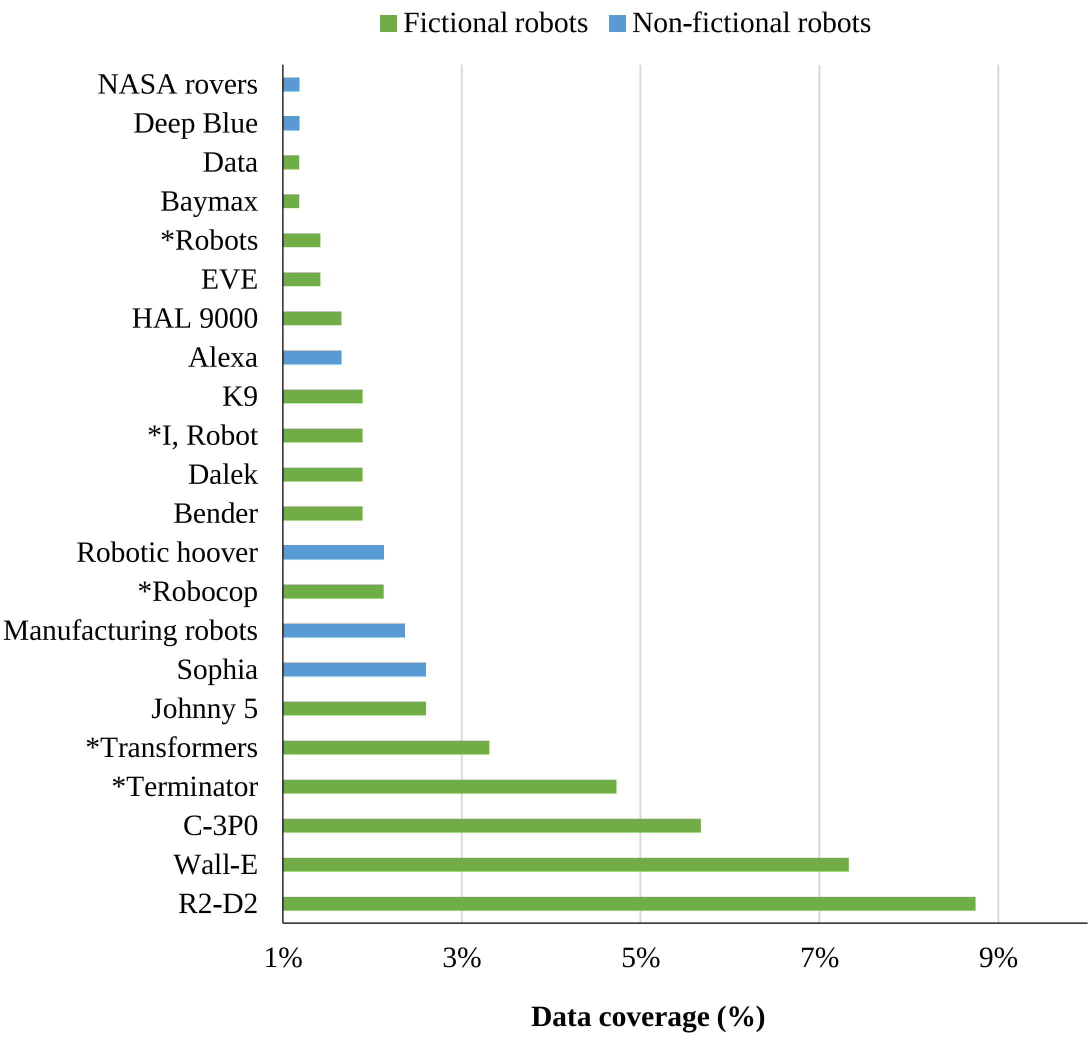
Salience of fictional vs. non-fictional robots. Participants mentioned a significantly greater number of fictional robots (n = 291) than non-fictional robots (n = 123), $\chi^2(1) = 68.17$, $p < .001$. This finding suggests that fictional representations of robots are more salient in the general population than non-fictional representations. In line with the finding that participants mentioned more fictional than non-fictional robots, the top five most mentioned entities by participants were all fictional (see Figure 3.5) with "R2-D2" from the Star Wars movie franchise being the most mentioned robot. The most mentioned non-fictional robot was "Sophia" (a realistic humanoid robot developed by Hanson Robotics); followed by non-specific manufacturing robots. A full list of all entities, the number of times they occurred, and their origin and fictional status can be found in Appendix I.

Relationship between the number of fictional robots and explicit and implicit attitudes. A regression was conducted to see whether the percentage of fictional robots that participants mentioned predicted each of the four dependant variables: attitudes toward interaction with robots (NARS-S1), attitudes toward the social influence of robots (NARS-S2), attitudes toward emotion when interacting with robots (NARS-S3); and implicit attitudes as measured via the IAT (d-scores). Statistical assumptions for regressions were checked following the same procedure detailed in Section 3.3.2.2.

The percentage of fictional robots that participants mentioned did not predict their attitudes toward interaction with robots (NARS-S1), $F(1, 98) = 0.02$, $p = .887$, $R^2 = 0.00$; attitudes toward the social influence of robots (NARS-S2), $F(1, 98) = 0.83$, $p = .365$, $R^2 = 0.01$; attitudes toward emotion when interacting with robots (NARS-S3), $F(1, 98) = 1.08$, $p = .302$, $R^2 = 0.01$; or their IAT scores, $F(1, 98) = 0.04$, $p = .848$, $R^2 = 0.00$.

Figure 3.5

Entities Mentioned by Participants that Occur in Over 1% of the Data



* The majority of entities mentioned under these categories were non-specific with a few exceptions where specific characters from the movies were mentioned by name. They were not counted separately.

3.3.3 Discussion

3.3.3.1 Semantic networks as a model of the social representation of robots

The main purpose of Study 3 was to gain insight into the way that people conceptualise *robots* by constructing a semantic network of word associations that taps into the associated social representations as defined by Abric (1993). In line with similar studies (Doerfel, 1998; Palla et al., 2005; Keczer et al., 2016), it was possible to construct such a network with features that supported Abric's proposed central and peripheral structure of social representations. More specifically, the constructed network was modular, consisting of a few global hubs (i.e., associations) that accounted for the majority of connections in the network and were more frequently observed than the majority of other associations. These properties were consistent with Abric's description of the central core which is both homogeneous (i.e., common) across individuals, defines the meaning of the social representation and underlies the more flexible peripheral system. Additionally, it was found that removal of such hubs would result in complete or partial dissolution of the network's modularity and connectivity. This is again in line with the core concept of social representations without which a socially-relevant category will have no coherent and consistent existence in a social group. The modular, scale-free nature of the network also meant that the majority of associations not only occurred much less frequently than the global hubs but were also less connected to other associations. This is again in line with the Abric's peripheral component of social representations which is more flexible and varied across individuals, and is not responsible for the underlying structure and organisations of the representation.

3.3.3.2 The core component of the social representation of robots

The core component is represented by all the word associations in the network that have an outstanding number of connections to other associations and occur considerably more often in the network (i.e., the global hubs). One hub in particular stands out as having such properties, namely "metal", and the findings suggest that this association is at the core of how people represent robots. According to the findings of a number of other studies (Wachelke & Lins, 2008; Dany et al. 2015; Keczer, et al., 2016) the central elements of a representation are often abstract concepts and characteristics (e.g., fear). This is clearly not the case for the social representation of robots as "metal" is a rather concrete property. This finding may suggest that when people are asked to think about robots they are more likely to retrieve an example of a metallic robot. According to Lord and Lepper's Attitude Representation Theory, this may have an impact on the stability of people's attitudes toward robots due to a potential mismatch between individuals' subjective representation and real robots (e.g., the Paro robot which has a fur cover). Exposure to real robots, especially where direct contact is a novel experience that contradicts participants' existing ideas about robots, may be particularly impactful in terms of attitudes. One possible explanation of why "metal" has such a central role in the social representation of robots may be for historical reasons. Both early and current robots are often entirely or partially made of metal which may be a particularly striking feature if the robot's appearance mimics that of a living entity such as a human. However, whether this is true or not is impossible to confirm without further investigation. Alternatively, it could be that robots are consistently depicted as metallic in fiction. There is some evidence for this connection in the findings for Study 3 as the most commonly mentioned fictional representations of robots by participants (i.e., R2-D2, Wall-E, and C-3P0) are all made of an apparently metallic material. Another notable property of this association is its neutral valence as only two participants stated that "metal" had a positive rather than neutral meaning to them. This indicates that while "metal" may be at the core of the social representation of robots, it does not necessarily contribute to a particularly positive or negative view of robots at a group level. Since we know that the evaluative elements of social representations are generally attributed as most influential in terms of attitudes (Moliner & Tifani, 1997), if the valence of the associations can be argued to be evaluative, then it may be an indicator of particular aspects of the representation that are influential in terms of people's attitudes (e.g., very positive or negative associations).

The core of the social representation found in Study 3 differs from the one that Piçarra et al. identified in their study. Specifically, while "metal" was still a relatively high-frequency association, it was located in the periphery of the representation. Instead, Piçarra et al. found that "machine" was the association that was both most frequently evoked and ranked highest thus was at the core of the social representation. In Study 3, "machine" was the third largest global hub but definitely not comparable to "metal" (although, strongly connected to it in the network as most global hubs are). In addition, according to Piçarra et al., the only notable connection that "machine" had was to "computer" and "technology", with "technology" becoming central to the representation despite it being previously located in the periphery. This is somewhat unexpected as, according to Abric, the core of any social representation is responsible for its organisation and structure and as such should be connected to many other elements of the representation as is the case with "metal" in Study 3. However, this peculiar finding by Piçarra et al. may be largely due to limitations with the way the data was analysed rather than a true reflection of the social representation (see Section 3.1.4).

Study 3 also found that the central component of the social representation comprised of associations including "AI/artificial intelligence", "machine", "mechanical", "helpful", "emotionless", "artificial", "technology", "computer", and "useful". Once again, the majority of these network hubs appeared to represent concrete, predominantly neutral (according to participants) properties rather than abstract concepts. In terms of valence, there were three notable exceptions, "useful" and "helpful" that were completely positive associations and "emotionless", a completely negative association. Although at the opposite ends of the spectrum, at least where valence is concerned, these three concepts were not necessarily contradictory (i.e., many things are useful, helpful, and emotionless). In fact, none of the associations found at the core represented conflicting ideas. This is consistent with Abric's description of the core of social representation as a "coherent" structure while the peripheral system allows for contradictions (e.g., that robots are both "dangerous" and "practical"). The consistency among core associations may indicate that, at a group level, robots are viewed as neutral and somewhat removed from the social domain. This may however be beneficial where robots are concerned as this neutral core may allow for greater flexibility of the peripheral structure of the representation which is formed as a response to immediate contexts within which the representation is relevant. Overall, the findings from Study 3 suggest that robots are generally viewed as artificial creations (i.e., machines) that, although apparently intelligent and useful, are emotionless.

3.3.3.3 Fiction as a part of the social representation

As discussed in Section 3.1, fiction has been suggested to play a large role in shaping the way people think about robots. If this is indeed the case, ideas related to fiction should be present in the social representation of robots. Such a presence is not directly apparent in the core of the social representation as none of the global hubs can be said to represent concepts that explicitly relate to fiction. However, Module 1 contained almost all associations that are directly related to fiction: "Asimov", "book", "cyborg", "fiction", "movies", and "sci-fi" (and possibly "android", although the use of the term is no longer constrained to fiction only). Out of all of these associations, “sci-fi” (i.e., science-fiction) was the only relatively large node comparable in size to “clever” and “help”, indicating its importance in the peripheral structure of the representation. More importantly, “sci-fi” was strongly connected to the core of the representation and the biggest modular hub (i.e., “metal”), providing some evidence toward the earlier assertion that fictional representations of robots may have contributed to the idea of robots as almost exclusively made of metal. Within Module 1, “sci-fi” was also strongly connected to another core elements of the representation, “artificial intelligence”. This finding further supports the idea that fiction is impactful in the way people conceptualise robots and possibly has some influence in shaping their attitudes toward robots. It should also be noted that the average valence of “sci-fi” is slightly positive and no participants indicated that the association held a negative meaning. This could mean that although robots in fiction are often depicted in an unrealistic way or, in many cases, as sinister characters, science-fiction may not be a key part of the representation of robots that contributed to negative attitudes. This is somewhat in line with Bruckenberger et al. (2013)’s findings that previous experience with fictional representations of robots can lead to variable attitudes when participants are confronted with indirect or direct contact with non-fictional robots and that any fictional representation (be it a portrayal of a ‘good’ or ‘evil’ robot) may lead to ambiguous attitudes that are not based on reality. Although Piçarra et al. did not focus on fiction, it should be noted that “movies” was one of the most frequently made associations in the periphery of the representation while “fiction” was also present but much less frequently occurring. This supports the findings from Study 3 and indicates that fiction is present in social representation of robots across time and different cultures. In addition, a number of references to fictional robots, “Robocop”, “I, Robot”, and “Star Wars”, also occurred although less frequently. Such direct references to fictional robots were not present in the network constructed in Study 3, although this may have been as a result of the first task participants in the study performed (i.e., listing the first five robots they could think of). Interestingly, all three of the references to fictional robots were in the top fifteen most mentioned robots in Study 3, providing support for the salience of specific fictional robots.

3.3.3.4 Modular organisation of the network

The semantic network was divided into five distinct modules representing different aspect of the social representation of robots. It can be assumed that elements (i.e., associations) in modules share similar properties (i.e., meaning) within the representation. The Louvian Community algorithm used in Study 3 to investigate the modularity of the network is essentially a more sophisticated way to divide the elements of the social representation into meaningful categories than the method used in Study 2 and the method used by Piçarra et al., both of which relied on the researcher's interpretation of the qualitative meaning of the data. A broken-down visualisation of the modules and the valence of the nodes in each module can be found in Appendix F.

Core aspects. Module 1 contained the largest number of global/modular hubs (i.e., "metal", "AI/artificial intelligence", and "computer") and is arguably the module that contained the majority of the core of the social representation. The three modular hubs were strongly interconnected (as is expected for the core elements of the representation) and accounted for the majority of links between associations, further establishing their role in determining the organisation of the module and the network as a whole. As mentioned, Module 1 contains almost all associations related to fiction as well as a few associations that reference various characteristics of robots (e.g., "funny", "silver", "metal", "wires"). These findings are in line with Piçarra et al.'s observations of the thematic link (i.e., "robot characteristics") between many of their high-frequency associations. The findings are also partially in line with the largest category established in Study 2, "Artificiality", that contained all associations relating to metal and artificial intelligence (see Table 3.2). However, it would be difficult to reduce Module 1 to simply a part of the representation that demonstrates the artificial nature of robots as there are a number of elements that do not necessarily fit this description. Of particular note is the fact that, overall, Module 1 contains largely neutral associations with only four elements ("movies", "beep", "help", "funny") that had notable positive meaning and only two associations ("apocalypse" and "cyborg") that had notable negative meaning. Out of these associations, only "help" was a relatively large node, indicating that the associations with a strong valence likely do not play a large role in the social representation. Overall, Module 1 appeared to contain a variety of ideas centred around the core of robots as artificial creations made of metal with obvious historic ties to the field of artificial intelligence and computing.

Potential of robots. In contrast to Module 1, Module 2 contained almost exclusively positive associations centred around two strongly connected global/modular hubs, "helpful" and "useful". This module appears to correspond to the idea of a "helpful machine" that Piçarra et al. identified as one of the themes among high-frequency associations. The ideal of a "helpful" robot was also strongly connected to "clever" (although "intelligent" is also located in this module), which likely reflects the common perception of robots as cognitively able (De Graaf & Allouch, 2013). In addition, Module 2 also contained associations (e.g., "advanced", "development", "futuristic") that could fall into one of the thematic categories identified in Study 2, namely "Technological advancement". "Technology" (part of the core of the representation) and "science" were also located in Module 2, adding to the idea of robots as an outcome of human progress. There was one notable excepion to the overreaching theme of Module 2, and that is the idea of robots as "scary" which was the only completely negative association in the module. "Scary" was strongly connected to two elements of the representation, "clever" and "intelligent". This finding might indicate that, although the potential of robots' cognitive abilities may be viewed as helpful, it may also present a certain level of threat. Despite this, Module 2 largely represented the idea of robots as part of progress and potentially helpful and useful tools in the future.

Emotionless. Module 3 is the only module that has a global/modular hub, namely "emotionless", that is negative. In this module, "emotionless" is strongly connected to "mechanical", the second global/modular hub in Module 3. This connection further emphasises that robots are, at the core, viewed as emotionless machines. Module 3 contained by far the most contradictory (both in meaning and valence) peripheral elements of the social representation (e.g., "cold", "slow", and "expensive" vs. "logical" and "practical"). Overall, Module 3 represented the conflicting qualities of robots that likely reflect the variety of fictional and non-fictional robots that people are exposed to.

Artificial constructs. Module 4 was a considerably smaller module with "artificial" as the single global/modular hub. The majority of associations were slightly to completely positive in valence with the only notable exception being "jerky (movements]" that fitted with the idea of artificiality but held a negative meaning. Overall, this module largely supported the idea at the core of the social representation of robots as "metallic" and "artificial". However, unlike Module 1, Module 4 framed artificiality more concretely in "engineering", "automation", and "electronics" to which "artificial" was strongly connected.

Machines. Module 5 is the only module that did not have a modular hub despite containing the global hub "machine". Although "machine" is neutral and the largest node in the module which accounts for the majority of connectivity, all other associations are either strongly negative (“uncanny”, “dangerous”, and “large”) or positive (“powerful” and “fast”). In Piçarra et al.’s paper, “machine” was at the core of the social representation with a contradictory periphery that, according to Piçarra et al., represented the various embodiments and roles robots could take (be it in fiction or reality). Module 5 supported the idea of a neutral concept of a machine to which a variety of contradictory properties could be applied to account for the various current and future roles of robots.

3.3.3.5 Salience of fictional and non-fictional representations of robots

Findings from Study 3 confirmed the expected discrepancy between the number of fictional and non-fictional robots that people spontaneously bring to mind. As hypothesised, fictional robots were more salient examples of the robot category than non-fictional ones. This is to be expected as the general public is much more likely to have come into contact with fictional representations of robots rather than real ones. This unbalanced experience with robots has been cited multiple times as a possible factor that affects people’s expectations of robots’ capabilities and thus their interaction with robotic devices (e.g., Broadbent et al., 2009; Kriz et al., 2010; Nomura et al., 2008, among others). Since some fictional robots appear to be frequently brought to mind (e.g., R2-D2, Wall-E, Terminator, etc.) and originate from popular and widely accessible media such as films, it is not unlikely that they have some role to play in the formation of the core of the social representation of robots. “Metal” has already been mentioned as one feature of the core to which consistent representation of robots in fiction may have contributed. Another notable association to which fictional robots may have contributed is “emotionless”. According to a content analyses of popular science fiction films by Keczer et al., robots are consistently depicted as having superior physical and cognitive abilities but lacking in social and emotional intelligence. Such depictions may have contributed to the “emotionless” property at the core of the social representation of robots. The association between robots and “emotionless” was qualitatively explored in Study 4 and offers some support for the suggested impact of fictional depictions.

Another interesting question regarding the salience of fictional and non-fictional examples of robots may be if and how prominent examples have changed and will change over time. Although, once again there is not much research that could be used to answer this question, the data from Study 3 indicates that although fictional robots are still at the forefront of people’s minds, some non-fictional robots are mentioned more frequently.

Namely, the humanoid robot Sophia that has made increasingly more appearances in non-fictional programs and news outlets for the past two years. This may indicate that popular non-fictional media could play a large role in introducing the general public to, hopefully, more realistic examples of robots than fictional media.

In terms of the relationship between fictional robots and people's attitudes, no evidence was found in Study 3 that there was a relationship between the number of fictional robots that people mentioned and their explicit and implicit attitudes. This is somewhat surprising as a study by Riek et al., (2011) found a positive correlation between the number of fictional films people have watched and their attitudes as measured using the NARS. However, as Riel et al. acknowledged, it is possible that people who are already interested in robotics and have more positive attitudes are also more likely to engage with fiction but may in turn also be more familiar with real robots. Additionally, being able to name fictional robots does not necessarily equate to engagement with fictional media as some robots (e.g., R2D2 from the Star Wars franchise) are pervasive in popular culture and may be known to multiple individuals. As no information regarding people's experiences with fictional robots was collected in Study 3, it is difficult to draw any conclusions with regards to the above mentioned finding.

3.3.3.6 Limitations of the study

One limitation of the study concerns the lack of ranking of the word associations which are generally considered to be hierarchical. Meaning that not only are the first associations considered stronger but also, to a certain extent, determine the subsequent associations (Dany et al., 2015). Due to methodological limitations, it was not possible to consider the rank of the associations. It is possible that such an omission may inflate the importance of some associations as presented in the network. For example, while "metal" is the most frequently made association it is still be possible that it is has a lower rank and is thus of less importance in the central system of the representation.

A second limitation relates to the construction of the network. As described by Steyvers and Tenenbaum, the scale-free, modular properties of networks are best detected for large networks (roughly defined as networks with more than 100 nodes). Although it can be said with reasonable certainty that the network constructed in Study 3 is scale-free, it only consists of 79 nodes. This could have impacted the partitioning of the network (i.e., the way the nodes are divided into individual modules) which in turn could have resulted in a less accurate model of the social representation of robots. Additionally, the Louvian Community algorithm that was used to determine the modularity of the network is also meant to be used with large data sets for optimal results (Blondel et al., 2008).

A third limitation is that Study 3 did not explicitly consider whether and how the nature of people's representations of robots shaped their attitudes. As such, Study 3 did not inform the relevance of Lord and Lepper's Attitude Representation Theory with regards to the variability in people's attitudes toward robots. The qualitative work presented in Study 4 addresses this limitation and further built upon the findings of Study 3.

3.3.3.7 Conclusion

Study 3 has expanded upon Piçarra et al. work on the social representation of robots as broad socially-relevant group by utilising semantic networks in order to visualise the core and peripheral structure of the representation. At its core, the social representation of robots is a neutral one, rooted in historic connections between the field of robotics, computing, artificial intelligence, and technology more broadly. The findings from Study 3 indicate that people view robots as essentially emotionless metal constructs but a periphery of often contradictory ideas and properties suggest that the variety of fictional and non-fictional robots as well as their potential usefulness is not missing from the conceptualisation of robots. Although not a part of the central representation, concepts related to fiction were connected to the core of the social representation, giving some support to the idea that fiction has shaped the way people think about robots. This is further supported by the relationship between the valence of associations within the network and participants' explicit attitudes, providing some evidence that social representations may impact attitudes. However, the meaning of the modules was largely open to interpretation, especially given the absence of literature on the topic. Study 4 did however support the interpreted meaning of the five modules as the themes that emerged from the qualitative analysis were largely synonymous with the conclusions drawn from Study 3. Ultimately, Study 3 was not appropriate for drawing any conclusions regarding the link between people's representations and attitudes. This topic was instead informed by the qualitative analysis of semi-structured interviews in Study 4 that focused on teasing apart the structure of individual representations with reference to the findings from Study 3.

3.4 Study 4 - Understanding the Basis of People's Attitudes Toward Robots

The findings of Study 3 supported the diversity of individuals' representations of robots which were divided into five distinct modules reflecting different aspects and characteristics of robots. Study 4 further explored these modules using a different method – namely, the thematic analysis of semi-structured interviews. This approach was chosen as it allowed for an in-depth exploration of the modules and meaning behind high-frequency associations that may otherwise not be possible by using purely quantitative methods. Furthermore, thematic analysis allowed for a more structured conceptualisation of participants' mental representation of robots based on the abstract definition provided by Lord and Lepper's Attitude Representation Theory. A number of steps were added to the traditional procedure of thematic analysis with the most notable being the coding of sentiment (*positive, neutral, negative*) which was compared to participants' reported explicit attitudes as measured by the NARS.

Study 4 also investigated to what extent each module was present in individuals' representations of robots and how the presence of each module shaped people's attitudes toward robots. Doing so allowed for evaluating the extent to which Attitude Representation Theory could be used to explain the variability of people's attitudes toward robots and provided further insight into what aspects of the social representation of robots were likely to influence people's attitudes.

Finally, Study 4 investigated the role of participants' experiences with fictional and non-fictional robots on the relationship between participants' representation of robots and their attitudes in order to inform prior claims regarding the importance of experience in shaping people's attitudes toward robots (Riek et al., 2011; Bruckenberger et al., 2013).

3.4.1 Method

3.4.1.1 Participants

Fifteen participants (9 females, 6 males) between the ages of 21 and 64 (M = 37.33, SD = 13.17) took part in a semi-structured online interview between June 2nd and June 23rd, 2020. Participants considered themselves fluent in English and were not professionally involved in robotics. Four of the participants indicated that they were *extremely interested* in robotics, six indicated that they were *somewhat interested*, one was *somewhat disinterested*, and one participant did not provide information.

Sample size was determined by practical considerations and the likelihood of reaching theoretical saturation based on Ando et al., (2014) who found that 92% of all codes emerge after inductive thematic analysis of data from the first 12 participants. Effort was made to recruit participants from the same population in Study 3 by adopting a similar recruitment strategy (see Section 3.3.1.1). More specifically, the study was advertised via mailing lists comprised of staff and student volunteers at the University of Sheffield and through a Facebook post which was shared by members of the researcher's personal social network.

3.4.1.2 Procedure and Materials

Pre-interview questionnaire. It was important to select a sample that was at least partially comparable to the samples requited for the other studies presented in this thesis. It was also necessary to ensure that participants had the necessary equipment to take part in the online interviews. As such, participants were first asked to complete an anonymous survey to check that they were eligible to take part via Qualtrics asking them the following questions: "Do you consider yourself fluent in English?", "Do you have access to a device with working camera and microphone?", and "Do you work or have you ever worked within the field of robotics?". Participants who met the eligibility criteria were forwarded to another Qualtrics survey containing the information sheet, consent form, timeslot booking for the interview, and the pre-interview questionnaire.

Interview timeslot booking. Participants were asked to provide their name, age, gender identity, and email address. They were also provided with a randomly generated Participant ID which was used to anonymize their data and to book a 1-hour timeslot using the online scheduling tool *youcanbook.me*. Participants were emailed a reminder, instructions, and a link to the online interview a day before their chosen timeslot.

Experience with fictional and non-fictional robots. After booking an interview timeslot, participants were asked to report how many robot-related experiences they have had during their lifetime or in the past year (see Table 3.8) on a 10-point scale from 0 experiences to 10 or more experiences. This measure was adapted from the Robot-related Experiences Questionnaire used by MacDorman et al., (2009).

Negative Attitudes toward Robots Scale (NARS). Participants were then asked to complete a modified version of the Negative Attitudes toward Robots Scale (NARS) to measure their explicit attitudes toward robots (Nomura et al., 2004). The NARS is comprised of three subscales (*Interaction with robots*, *Social influence of robots*, and *Emotion in interaction with robots*) with a total of 16 items (see Table 3.4, Section 3.3.1.2). Each item is a statement that can be rated on a five-point scale from 1 (strongly disagree) to 5 (*strongly agree*). The order in which the items were presented was randomised. The reliability and validity of the NARS has been supported by multiple studies (Nomura et al., 2004; Nomura et al., 2006). Participants' mean for each of the three subscales of the NARS was generated by summing up the score for each of the items in the three subscales as per the instructions in Nomura et al. (2004). The scores of the NARS-S3 subscale were inverted prior to summation. For the NARS-S1 subscale, the minimum possible score is 6, the maximum score is 30, and the mid-point is 18 (indicating neutral attitudes); for the NARS-S2 and NARS-S3 subscales the minimum is 5, the maximum is 25, and the neutral mid-point is 15. Larger values indicate more negative attitudes, while smaller values indicate more positive attitudes. For each type of experience with robots, the number of experiences that each participant reported were summed so that the resulting value was between 0 – 20 for indirect experiences with non-fictional robots and fictional robots, and between 0 – 30 for direct experiences with robots.

Table 3.8

Robot-related Experiences Questionnaire

Type of experience	Item
Experiences with fictional robots	Read stories, comics, or other fictional material about robots. [a]
	Watched movies, fictional TV programmes, or other media about robots. [a]
Experiences with non-fictional robots (indirect)	Read news articles, product descriptions, conference papers, journal papers, or other factual material about robots. [a]
	Watched documentaries, factual TV programmes, or other factual media about robots. [a]
Experiences with non-fictional robots (direct)	Built or programmed a robot. [b]
	Had physical contact with a robot (not including appliances such as robotic vacuum cleaners). [b]
	Attended lectures, exhibitions, trade shows, competitions, or other events related to robots. [b]

[a] Number of activities performed in the past 1 year.

[b] Number of activities performed in one's life so far.

3.4.1.2.1 Interview guide

The interview guide (see Table 3.9) was constructed in line with the aims of the study, following the recommendations for semi-structured interviews detailed in Newcomer et al. (2015). Questions 8-17 were the main focus of this study, with one question assessing the extent to which participants associated robots with the representations reflected by each of the modules identified in Study 3 (see Section 3.3.2) and another question assessing whether participants believed that representation affected their attitudes. One or two associations were selected for each module based on their intramodular node strength and how well they were though to represent each module. For example, the word "artificial" was selected to represent *Module 3: Artificial constructs* as it was the largest node for that module and was representative of the overall modular theme. The order of the questions in Table 3.9 was mostly preserved but varied depending on participants' responses. In general, questions 1-7 were addressed first and acted as an icebreaker that provided a natural escalation to the more in-depth portion of the interview.

Table 3.9

List of Interview Questions

Code	Question
Q1	What image pops into your head when you think of "robots"?
Q2	Why do you think that image came to mind?
Q3	Do you think that other people would imagine something similar? Why?
Q4	Do you think that what you imagined is representative of robots as a whole?
Q5	Can you give me a couple of examples of fictional robots?
Q6	Can you give me a couple of examples of real robots?
Q7	In what ways are fictional and real robots similar and/or different?
Q8	Some people associate robots with the words "computing" and "artificial intelligence". Do you associate robots with "computing" and "artificial intelligence"?
Q9	Do you think that associating robots with the words "computing" and "artificial intelligence" affects your opinion on robots? Why do you think that is?
Q10	Some people associate robots with the words "helpful" and "useful". Do you associate robots with "helpful" and "useful"?
Q11	Do you think that associating robots with the words "helpful" and "useful" affects your opinion on robots? Why do you think that is?
Q12	Some people associate robots with the words "emotionless" and "cold". Do you associate robots with "emotionless" and "cold"?
Q13	Do you think that associating robots with the words "emotionless" and "cold" affects your opinion on robots? Why do you think that is?
Q14	Some people associate robots with the word "artificial". Do you associate robots with "artificial"?
Q15	Do you think that associating robots with the word "artificial" affects your opinion on robots? Why do you think that is?
Q16	Some people associate robots with the word "machines". Do you associate robots with "machines"?
Q17	Do you think that associating robots with "machines" affects your opinion on robots? Why do you think that is?

3.4.1.2.2 Interview process

Setting. The semi-structured interviews were carried out using the video conferencing application Google Meet and took approximately 30 minutes and up to 50 minutes in one case. The researcher was always situated in the same quiet, indoor location with a plain wall as a background. Participants' environments, devices, and internet connection varied but most participants were at quiet indoor locations and no major technical difficulties occurred during the interviews.

Recording and quality of the interviews. The video and audio feedback from each interview was recorded using the inbuilt recording function of the Google Meet application and was stored automatically as an MP4 file in the researcher's Google Drive. No physical notes were taken during the interview but general impressions were noted following each interview and were used to aid the analysis. The overall audio quality of the interviews was good, although occasionally noisy due to participants' internet connection and environment. As such, it was not possible to transcribe all of the interviews with 100% accuracy and some of what participants said was inaudible. However, the inaudible portions of the interviews made up less than 1% of the total data set and, as such, did not interfere with the analysis of the data.

Structure and tone of the interviews. Prior to starting the interview, introductions were made and participants were reminded of the aim of the research project, their right to withdraw, and the overall interview process and duration. Once participants consented to be recorded, the interview began. A casual, conversational, and friendly but professional approach to conducting the interviews was taken in line with recommendations from Newcomer et al. (2009) and Salmons (2014). The questions in the interview guide were generally asked in the order presented but further probes and / or additional questions were asked where appropriate or where it was thought that they would provide a better understanding of participants' experiences and views. Personal or sensitive questions were avoided. After all of the questions were asked, participants were given the opportunity to ask questions or to discuss anything that they wanted to in an informal conversation. This part of the interview was not transcribed or analysed. Finally, participants were debriefed verbally (see Appendix D for debrief procedure).

3.4.1.2.3 Transcription of the interviews

All audio files and transcriptions were anonymized using randomly generated Participant IDs and participants were assigned single letter pseudonyms that are used throughout this report. As the researcher transcribed the interviews, the transcription process was not anonymous. In one participants' case, it was necessary to make small changes to grammar and sentence structure in order to improve the readability of the script, although this was done tentatively and only where absolutely necessary. Personal information was omitted from the transcriptions but nonverbal features (e.g., a participant receiving a call) were noted. Filler words (e.g., "like"), repetitions (e.g., "I think, I think"), and sudden breaks in sentences were transcribed but mostly omitted in direct quotes in order to improve readability. However, filler sounds (e.g., "erm") and pauses were not transcribed. As such, the transcription could be described as primarily non-verbatim.

3.4.1.2.4 Coding and data analysis

The data was analysed using thematic analysis in line with recommendations from Braun and Clarke (2006; 2012) using the data analysis software NVivo 12. Braun and Clarke recommend six stages to thematic analysis; to which two additional stages were added to answer the specific questions posed by the research. First, in order to gain familiarity with the data, post-interview notes and transcripts were re-read and general observations about what was said were noted down.

Next, an additional step was taken to identify content that was potentially relevant to each of the modules identified in Study 3 (see Section 3.3.2) by conducting a word search of all associations (and stemmed words) related to each of the five modules (see Appendix G). Search results were reviewed and content that was consistent with the meaning of the modules was coded as belonging to that module (see Table 3.10). Next, content relevant to each of the 17 questions (see Table 3.9) was identified and coded as belonging to each question. This chapter only reports the findings relating to questions 8-17 and as such the analysis described in this section is relevant only to the content specific to those questions.[7]

---

[7] The analysis was constrained in such a way in order to ensure the coherency of Chapter 3 and ensure that the thesis length was manageable.

Table 3.10

Proportion of Data Coded as Relevant to Each Module

	Data coverage
Module 1: Core aspects	37%
Module 2: Potential of robots	46%
Module 3: Emotionless	25%
Module 4: Artificial constructs	21%
Module 5: Machines	21%

The second stage of the analysis generated the initial codes. Data from all participants relating to each question was coded using a mixture of descriptive (e.g., "humans must oversee robots") and interpretive codes (e.g., "fear of autonomous robots"). After all of the data was coded, the codes were revisited and, where appropriate, merged or re-worded to better capture the meaning of the data.

The third stage of the analysis searched for themes. Specifically, the data relevant to each module was reviewed together with the codes, and broader patterns of meaning were identified and labelled. All codes were assigned under initial themes which were then merged, reassigned, or discarded in a cyclical process of reviewing the data and codes.

The resulting list for each module was then reviewed in the fourth stage of data analysis where multiple themes were collapsed into single ones or entirely discarded if they were not supported by the coding or were too *thin* (i.e., lack data supporting the theme). This process was repeated multiple times until the themes were though to sufficiently cover the meaning of the data.

In the fifth stage of the data analysis, the themes were named and briefly defined. The themes were also ordered in such a way as to reflect the narrative of the data and the most prevalent themes across the participants were generally placed first as they typically addressed whether participants associated robots with module-specific words or not. At this stage, relevant quotes were extracted for each module in preparation for writing up the results.

Prior to the final stage of the analysis, an additional step was taken to code the sentiment (positive, negative, and neutral) for each interview segment and whether there was evidence of participants referring to experiences relating to fictional and non-fictional robots in their answers. Sentiment was coded by looking at the data relevant for each theme as well as the surrounding data and deciding whether the participant said something positive, negative, or neutral about robots. Each of the three valences was assigned a percentage such that if what a participant said regarding a topic was entirely positive, the segment was assigned a value of 100% for positive, and 0% for negative and neutral. Similarly, if what a participant conveyed about robots was both positive and negative in equal measure, that segment was assigned a value of 50% for positive and negative, and 0% for neutral. Additionally, each segment and the surrounding data were reviewed and given a value of 1 (*present*) or 0 (*not present*) for whether it contained references to participants' experiences with fictional and non-fictional robots. The resulting coding was visualised in an excel sheet and aided the final stage of the analysis. A summary of the average sentiment and experiences can be found in Table 3.13.

The sixth and final stage involved identifying and writing up how each theme connected to other themes for the module and how each theme related to the attitudes and experiences participants reported in the pre-interview questionnaire.

3.4.1.2.5 Statistical analysis

As a final step of the data analysis, post-hoc *t*-tests were conducted to compare the NARS scores of participants who either endorsed or did not endorse each association. Whether a participant endorsed a particular association or not was determined by their answer to the following questions: Q8, Q10, Q12, Q14, and Q16 (see Table 3.9). It should be noted that this was not an attempt to apply a statistical analysis to a qualitative data set but rather a quantification of *yes-no* questions which were then entered into statistical analysis alongside quantitative survey data. Although unusual, the application of a statistical analysis in this case was intended as supplementary to the main qualitative analysis and not a substitution of it. It was conducted in order to simplify the relationship between the endorsement of associations (indicative of the way participants represented *robots*) and participants' attitudes which was identified using thematic analysis.

Methods of quantifying and analysing qualitative data using specialised statistical analyses do exist and were briefly considered. However, such methods are typically used and appropriate for large data sets where a qualitative analysis is not feasible (Schwartz & Ungar, 2015). For example, qualitative data obtained from social media platforms (e.g., Twitter) via text-mining techniques and tools can be analysed using Automated Content Analysis (ACA) which is used to identify themes in a given data set, their frequency, and the relationship between themes by utilising algorithms and, more recently, machine learning (Boumans & Trilling, 2016).

3.4.2 Results

3.4.2.1 Pre-interview questionnaire

Table 3.11 shows that on average, participants reported positive attitudes toward interaction with robots (NARS-S1) and emotion during interaction (NARS-S3) but neutral attitudes toward the social influence of robots (NARS-S2). NARS-S3 was the least variable subscale with a single participant reporting somewhat negative attitudes. The first and second subscales were more varied but only for the NARS-S2 did more than one participant score more than two points above the neutral mid-point (thus indicating negative attitudes). As such, the second subscale was most useful when teasing apart the relationship between people's representation of robots and their attitudes. In terms of the experiences participants reported, only direct experiences with robots had a limited range with $N = 6$ participants reporting no direct experiences.

Table 3.11

Descriptive Statistics for the Pre-Interview Questionnaire

	M	*SD*	Range
Recent experiences with fictional robots	6.86	6.44	0 - 20
Recent indirect experiences with non-fictional robots	7.86	7.78	1 - 30
Direct experiences with non-fictional robots	3.29	4.81	0 - 14
NARS-S1: Interaction with robots	14.64	4.25	8 - 21
NARS-S2: Social influence of robots	15.14	3.53	9 - 22
NARS-S3: Emotion in interaction with robots	12.21	2.78	6 - 17

Note. All but one participant completed the NARS and Robot-related Experiences Questionnaire and as such the descriptive statistics compiled in this table are based on 14 rather than 15 participants.

3.4.2.2 Themes

A breakdown of the associations that participants endorsed is provided in Table 3.12. In order to summarise simplify the findings of the qualitative analysis, post-hoc *t*-tests were conducted to compare the attitudes of participants who endorsed or did not endorse the associations (see Table 3.13[8]). Due to the fact that nearly all participants endorsed the association which represented the first two modules, this was only possible for the associations relating to Modules 3, 4, and 5. Unsurprisingly given the small sample size, only one of the *t*-tests was statistically significant. Participants who endorsed the association with “emotionless” and “cold”, on average reported significantly more negative attitudes toward the emotional impact of interacting with robots (NARS-S3) than participants who did not endorse the association (see Table 3.13). A similar pattern of responses was observed for the other two NARS subscales in relation to the association with “emotionless” but they did not reach statistical significance.

Eighteen themes were identified from the data relating to questions 8-17 with an average of 4 themes per module. Table 3.14 contains a summary of the themes, and each module is discussed separately in the proceeding sections.

Table 3.12

Break Down of the Associations that Participants Endorsed

	N	
Module	Endorsed	Did not endorse
1: "artificial intelligence" and "computing"	14	0
2: "useful" and "helpful"	12	2
3: "emotionless" and "cold"	9	5
4: "artificial"	7	7
5: "machines"	11	3

Note. One participant was excluded as they did not answer the pre-interview questionnaire.

[8] The findings presented in Table 3.13 are intended as an aid to interpreting the qualitative analysis and caution should be applied when interpreting their significance due to the sampling limitations.

Table 3.13

Descriptive Statistics and Independent t-test for NARS Based on Endorsement of Associations

Measure	Endorsed		Did not endorse				
	M	*SD*	*M*	*SD*	*t*(12)	*p*	*d*
Module 1							
NARS-S1	14.64	4.25	-	-	-	-	-
NARS-S2	15.14	3.53	-	-	-	-	-
NARS-S3	12.21	2.78	-	-	-	-	-
Module 2							
NARS-S1	14.25	4.48	17.00	1.41	-	-	-
NARS-S2	14.50	3.15	19.00	4.24	-	-	-
NARS-S3	12.00	2.95	13.50	0.71	-	-	-
Module 3							
NARS-S1	15.67	3.12	12.80	5.72	-1.04	.344	0.62
NARS-S2	16.22	3.49	13.20	2.95	-1.63	.129	0.93
NARS-S3	13.56	1.74	9.80	2.78	-3.14	.008*	1.62
Module 4							
NARS-S1	14.14	3.85	15.14	4.88	0.43	.678	0.23
NARS-S2	15.14	2.55	15.14	4.53	0.00	1.000	0.00
NARS-S3	12.57	2.51	11.86	3.19	-0.47	.649	0.25
Module 5							
NARS-S1	15.18	4.31	12.67	4.16	-0.90	.385	0.60
NARS-S2	15.55	3.36	13.67	4.51	-0.81	.435	0.47
NARS-S3	12.36	1.96	11.67	5.51	-0.37	.716	0.17

Note. Larger NARS scores denote more negative attitudes. *d* denotes Cohen's D.

Table 3.14

Summary of the themes identified with respect to each module and their prevalence (N); the average positive (+ve), neutral (nu), and negative (-ve) sentiment expressed in each theme; and the number of participants (%) in each theme who referred to experiences with fictional robots (FE) and experiences with nonfictional robots (NF)

Theme	Sentiment (%)			*N* (%)		
	+ve	nu	-ve	FE	NF	Example quotes
Module 1 – “artificial intelligence” and “computing”	28	49	24	19	75	
1: Artificial intelligence and computing are an integral part of robotics *N* = 9	0	100	0	0	100	“Perhaps like because whenever you say artificial intelligence, oh, you know, there's robotics.” “Robots need programming and computing …”
2: Concerns and consequences regarding technology advancement and use *N* = 7	13	23	64	57	57	“… it's a case of how far do we let AI go before they take over? … how far do we allow AI to take over before we then become extinct?” “… if they could create such a robot then the humans, right now they have the, the problem finding a job but it would be much worse if the scientists could create such a robot.”
3: Artificial intelligence contributes to the usefulness of robots *N* = 6	70	23	7	0	67	“… it [AI] advances the idea of robots that can be used just for specific things. For doing operations, or doing engineering things, industrial ones, so [they] could be more advanced with this new technology.” “It's not human so it can obviously do things like run numbers really quickly and it's quite incredible what technology can do. So, it is like artificial intelligence.”

Table 3.14 (continued)

Theme	Sentiment (%)			N (%)		
	+ve	nu	-ve	FE	NF	Example quotes
Module 2 – "useful" and "helpful"	80	3	17	5	67	
1: Robots are useful and helpful $N = 13$	92	0	8	8	85	"I think [having robots] it's a great thing. I think if we're able to do that, I don't know, it could help a lot of people." "I'd say helpful and useful, like I said, in the broad spectrum of things, say from the very technical application to that … kind of robot cat for a lonely old person or whatever. Or something that is you know more fun, educational … I'd say that's helpful."
2: Concerns and consequences regarding the use of robots $N = 8$	39	11	50	13	63	"… people might be wary of this kind of idea that a robot can take over human jobs and, you know. There won't be as many jobs for people to go around because robots will [take] them." "We've become so, reliant on it [technology]. That it's a little bit kind of it's a little bit dangerous, actually, too. …Getting too submerged and depending too heavily on technology and robots is probably not the best either."
3: Responsible use of robots $N = 5$	100	0	0	0	40	"It depends how we use them, you know. How we use these opportunities, how we use this instrument, devices. I think it's totally dependent on humans." "I think you'll always have them [robots] overseen. Checking. …Everything's different, it throws different things every time. Somebody overseeing. You'd always need that element."

Table 3.14 (continued)

Theme	Sentiment (%)			*N* (%)		
	+ve	nu	-ve	FE	NF	Example quotes
4: Envisioning how robots can be useful in the future *N* = 5	90	0	10	0	80	"I definitely see potential for things like cleaners, like I'm sure, in 20 years' time, like some things like cleaning could be done by robots. Basic manual labour. I can see them like building things." "Definitely a thing I could see working. Big lorry comes round, drops stuff off and off they go in an area and come back off they go again, I could see that sort of thing working."
Module 3 – "emotionless" and "cold"	22	43	35	16	43	
1: Current state of robots *N* = 8	19	75	6	13	100	"But machines, robots that we use in present are very, cold and clinical. Because they have a job to do and that's just their job." "No particularly strong feelings about it is just the sort of the level we're at technology wise at the moment. …we don't have a full AI really yet that can process and understand in that sort of on that sort of like human level yet."
2: Robots not capable of having real emotions *N* = 6	33	58	8	0	50	"In some ways, it's an artificial emotion that the robot is presenting, and that is very different from a real emotion." "I think that the missing part …would be that, it [emotions] will not be original because that's more of a human thing. …feeling is or emotions is the most human thing."

Table 3.14 (continued)

Theme	Sentiment (%)			N (%)		
	+ve	nu	-ve	FE	NF	Example quotes
3: Envisioning a future with emotional robots *N* = 9	36	37	28	11	0	"I think, yeah it would be positive (to have robots with emotions]. And I think also if we could, if it was possible to figure that out, it would be. In a lot of ways (robots will] be more intelligent than us" "If we got to the level of sort of something in like 'I, Robot' (a film], where you could have a friend who's a robot and it's not like you just like this robot because it's cool. It's like you've generally got an emotional human connection to it. That would be amazing."
4: Concerns and consequences of having robots with emotions *N* = 5	2	0	98	40	20	"Because if I imagine them being emotional, I just feel like it's crossing the line and crossing the boundary of how I can control it." "Well, I don't think you would really want an emotional robot to be honest. It's going to work and it doesn't fancy getting to work because it feels a bit down. That's not really good for your work, is it, really?"
Module 4 – "artificial"	18	57	25	38	22	
1: Defining the nature of robots *N* = 9	0	88	12	11	0	"I think that's [a] fair [association] because they are artificial, aren't they, really? I mean, they're made by somebody." "If we just consider artificial it means someone's, not some animals or whatever, but someone [human] has just created that, yeah [I would associate robots with artificial]."

Table 3.14 (continued)

Theme	Sentiment (%)			N (%)		
	+ve	nu	-ve	FE	NF	Example quotes
2: Robots are artificial humans *N* = 6	0	97	3	50	17	"You know, if I really associate it with artificial, it will be that this reason that they are like, fake versions (of humans], you know." "… but because of more my experiences and things like, as I said, with playing 'Fallout' (a game] with the synths, they are artificial humans with artificial intelligence. So, yeah, it's definitely an association I would make."
3: Concerns and consequences regarding technology advancement and use *N* = 5	18	0	82	40	20	"I think it potentially could be a little dangerous (to have robots] just because humans are fallible. So, if they're creating something there's potential to mess up. Essentially, make a mistake." "But if there is no balance and it's like all controlled by robots, or. It kind of doesn't make any sense. I mean, then it's what would humans do? Or you know, we can get lazy."
4: Artificiality of robots is a good thing *N* = 4	53	45	3	50	50	"… for me artificial is again related to being non-emotional. So, again if a robot with emotions, that will be a hard thing to deal with. So I think they are artificial." "I think in a way it's kind of a good thing (for robots to be artificial], because I think if robots were, you know, sentient, you know. Or like if they could think and feel, I think people would …get too attached to them."

Table 3.14 (continued)

Theme	Sentiment (%)			N (%)		
	+ve	nu	-ve	FE	NF	Example quotes
Module 5 – "machines"	17	60	23	13	67	
1: Defining the nature of robots *N* = 12	13	73	14	25	67	"I think every machine is a robot as well. Because it follows your command." "I suppose, yeah, to some degree (robots are machines]. To some degree. I mean, in the sense that they're unstoppable. Like once you turn a machine on, it can keep clonking away."
2: Robots distinct from machines *N* = 7	20	73	14	14	86	"Robots with a singular purpose seem to become more machine than robot in my mind, I think. ... Whereas if it's multi-purpose, it's more of a robot." "Machine could be mechanic or electric but with robots it's mechatronic. … So you can, you can upload, download, add the new applications so it's (robots are] more than just machines, or just, again mechanical or electronic thing"
3: Acceptance of robots as machines differs *N* = 6	18	32	50	0	50	"I think I would prefer robots to machines. I think they're more useful." "I guess for me it [machines] is not a positive term. Because it kind of makes you think a machine is something that doesn't have like a consciousness."

3.4.2.2.1 Emergent themes from the association of robots with "artificial intelligence" and "computing"

Artificial intelligence and computing are an integral part of robotics. This was the most prevalent (N = 9) of the three themes and it captured the historical and current relationship between robotics and the fields of computing and AI; as well as the integral part computing plays in how robots work. Eight participants identified AI and / or computing as being in some way inherently linked to robots immediately after being asked about the association. As Participants Q and E, respectively, stated "… whenever you say artificial intelligence, oh, you know, there's robotics" and "… (computing and robots] come hand-in-hand". All nine participants demonstrated or referred to some basic knowledge about the technical aspects of robotics such as Participant D who simply stated that "… there'll always be an element of, kind of, programming of some description to make a robot work". This is in line with participants' reported interest in and recent non-fictional experiences with robotics. All of the data for this theme was coded as neutral in valence (see Table 3.14) indicating that the initial link between robotics and AI and/or computing was descriptive in nature and not immediately invocative of any particular attitudes. As Participant D sated, "…it doesn't have a, kind of, negative connotation.". It seems unlikely that simply acknowledging the link between robots and AI and / or computing is particularly influential when it comes to people's attitudes toward robots and no discernible pattern was found in terms of participants' reported NARS scores.

Concerns and consequences regarding technology advancement and use. The second most prevalent theme (N = 7) encapsulates participants' concerns regarding the use of artificial intelligence (but not computing) in the context of robots and the consequences participants could envision as a result of utilising AI. However, on average, the sentiment participants expressed in relation to said concerns and consequences was not entirely negative in terms of how that made them feel toward robots (see Table 3.14).

For example, Participant A stated:

> Because like I said earlier, it's a case of how far do we let AI go before they take over? And it's the sci-fi part of my brain [that] thinks, how far do we allow an AI to take over before we then become extinct?

Although a seemingly negative take on AI, upon further probing of whether the association affected the participant's attitudes toward robots, Participant A expressed that they "… don't think what the general public would use [i.e., robots] are sophisticated enough to warrant any real concerns. … So no. Not everyday robots … I wouldn't be scared of them.". Participant A's response was consistent with their overall positive attitudes for all three NARS subscales. Participant C was similarly concerned despite their overall strongly positive attitudes, first saying that "[AI is] the scary part. That's probably more dangerous than it isn't." but then expressing a rather neutral, knowledge-based response to the association, consistent with their reported experiences with real robots: "See, the robotic technology, mechanical wise, is nowhere near as advanced as the software artificial intelligence.". Both Participant A and C reported multiple direct experiences with robots. This indicates that although someone might have legitimate concerns regarding AI, it may not necessarily impact their attitudes toward robots if participants have awareness and knowledge relating to the current state of robotics and the role of artificial intelligence.

Experiences with fictional robots may also play a role in people's perception of AI and its role in robotics as more than half of the participants (see Table 3.14) referred to fictional portrayals of robots in relation to their concerns. For example, going back to Participant A's initial response to the association, we can see that it is "… the sci-fi part of my brain [that] thinks …" about the consequences of utilising AI. This is consistent with the number of recent experiences with fictional robots that Participant A reported ($n = 7$) which is moderate when compared to other participants. Similarly, when Participant N was asked about what the connection between artificial intelligence and robots was, they said they're "… just going off of the fictional TV shows again, like 'Humans'." Unsurprisingly, Participant N reported some recent experiences with fictional robots ($n = 5$) but no direct experiences with real robots and had overall neutral attitudes toward robots. This difference between participants may indicate that although fictional portrayals of robots may raise various concerns about the role of AI in robotics, they may not necessarily affect attitudes if people have considerable real-life knowledge and experience with robots.

Artificial intelligence contributes to the usefulness of robots. This theme was the least prevalent of the three themes ($N = 6$) and captured participants' viewpoints regarding the beneficial role (e.g., "… (robots] could be more advanced with this new technology.") or potential of AI in relation to robotics (e.g., "… it can teach us things about ourselves."). Out of the six participants, three spoke about both the potential consequences and potential benefits of AI in the context of robots, and the other three participants only spoke about the potential benefits of AI; but there were no notable differences between those participants apart from the previously mentioned relationship between fictional experiences and concerns regarding AI.

On average, all participants expressed predominantly positive sentiment (see Table 3.14) when asked about the potential usefulness of AI in relation to robots. Participant K was one participant who did not explicitly indicate any concerns regarding artificial intelligence. When asked about the association, Participant K linked AI to the use of robots in a variety of fields, "… lots of things in the world … have robots to do things even now, you know, like on the space missions… Well, lots of areas really.", thus demonstrating some knowledge about real robots consistent with their reported number of experiences relating to non-fictional robots (n = 5). When probed about the effect of this association on their feelings about robots, Participant K stated:

> I think it (making the association] would veer me more towards the useful side of things. … Yeah, I think probably robots are in a lot more places than most people realize, to be honest. I mean, many people have them in their houses, don't they? They have those little robotic vacuum cleaners.

Although Participant K had overall neutral attitudes toward robots (as measured by the NARS), they did *disagree* with the NARS statement, *I would hate the idea that robots or artificial intelligences were making judgements about things,* which was consistent with their response. A similar pattern was found for three other participants.

Unlike the previous theme, *Concerns and consequences*, there was no evidence that participants' experiences with fictional robots informed their perception of the potential benefits of AI in relation to robotics. In fact, most participants referred to some knowledge or experience with robots and / or AI that was consistent with the number of recent experiences they reported relating to non-fictional robots. Overall, it appears that while the association between AI and robots may invoke a positive response regarding the usefulness of AI, this is not necessarily enough to affect participants' overall attitudes as measured by the NARS and other themes likely play a significant role.

3.4.2.2.2 Emergent themes from the association of robots with "useful" and "helpful"

Robots are useful and helpful. This was by far the most prevalent theme across all associations with nearly all participants (*N* = 13) recognising the potential usefulness of robots. Unsurprisingly, 12 of these participants indicated that they would associate robots with useful and helpful. Within this theme, three sub-themes were identified: the majority of participants (*N* = 10) spoke about *specific current and potential uses of robots* (e.g., "… a robot cat for a lonely old person"); some participants (*N* = 4) showed acceptance of the use of robots due to their *potential to be helpful in the future* (e.g., "… if it's going to aid what somebody is doing, it can only be a positive."); and some participants (*N* = 5) defined the usefulness of robots in terms of their *ability to do some jobs better than humans* (e.g., "… they can do that (making cars] a lot faster and a lot better than humans.").

On average, all participants expressed predominantly positive sentiment (see Table 3.14) in regards to robots being *helpful* and *useful*, although four participants' briefly mentioned some negative aspects like the potential of robots taking away jobs but the overall sentiment was not changed. As Participant D stated, "… it doesn't concern me personally but I can see why some people might, you know, might feel a bit threatened.". Nearly all (*N* = 11, see Table 3.14) participants directly referred to their knowledge and / or experience with real robots when speaking about their usefulness. For example, Participant F stated that "in the real world of seeing robots, [they've] been nothing but helpful and useful. … [In] things like manufacturing, they're great because they've improved cost, safety, time, everything like that …".

Overall, there are no discernible relationship between this theme and people's attitudes (as measured by the NARS) as nearly all participants associated robots with being helpful and useful. However, it is clear that most people think that robots are currently and / or potentially useful and their indirect and direct knowledge and experience with non-fictional robots clearly informs their opinions. Further exploration of what the two participants who did not associate robots with being *useful* and *helpful* may provide more clarity, as discussed in the next theme.

Concerns and consequences regarding the use of robots. This was the second most prevalent theme (*N* = 8) and it encapsulated participants' concerns regarding the use of robots (e.g., "… [robots] are also automating people out of jobs."); and the consequences they could envision as a result of utilising robotics in the future (e.g., "… depending too heavily on technology and robots is probably not the best."). This theme was closely interwoven with *Robots are useful and helpful* and eight of the participants mentioned both positive and negative consequences of using robots. Although the sentiment participants expressed was on average predominantly negative (see Table 3.14), seven of the participant expressed somewhat contradictory opinions, in varying degrees, about the consequences of using robots.

For example, Participant D stated the following:

> I think maybe people, people might be wary of this kind of idea that a robot can take over human jobs and, you know, there won't be as many jobs for people to go around because robots will take them. I suppose the counterargument is that then frees people up to do other things and it might save people, you know. Work places are safer and there's probably not as many accidents. So there's kind of there's a negative but then there's hopefully positives that come from it.

Even Participant H who stated that they would not associate robots with being helpful and useful, expressed that robots "… have the both sides. Negative and positive.", although this participant did not provide examples of the "positive" sides nor did they reference any experiences with fictional or non-fictional robots as a reason for their opinion. Overall, participants who expressed concerns regarding the use of robots had a varying number of experiences with fictional and non-fictional robots, and included participants who had overall negative, positive, and neutral attitudes as indicated by the NARS. However, the three participants who expressed concerns as a part of an overall positive or neutral narrative about the usefulness of robots had considerably more positive overall attitudes than other participants and were the only participants in this theme to *strongly disagree* with the NARS statement, *I would feel uneasy if I was given a job where I had to use robots.* The opposite was true for the participants who expressed predominately negative sentiment in relation to their concerns.

Similar to the findings for the second theme relating to the association with artificial intelligence, these findings show that being able to identify potential negative consequences of utilising robots does not necessarily mean that someone has negative attitudes. Likewise, not everyone who reported negative attitudes (as measured by the

NARS) expressed concerns when asked about the association between robots and *useful* and *helpful*.

Responsible use of robots. This was the third most prevalent theme (N = 5) that captured participants' views on how robots should be used. Participants explicitly acknowledged that while robots are or could be useful, they should be used responsibly. For example, Participant G said that "… it's about the responsible use of technology. Helpful. Useful. Yeah. Yeah, of course it is. But I feel that [we] need to create a balance as well.". This theme co-occurred with either or both of the previous two themes, further emphasising a somewhat double-minded view of the usefulness of robots. All five participants expressed entirely positive sentiment (see Table 3.14) regarding the responsible use of robots as it could improve the usefulness of robots and diminish some of the concerns participants expressed. Some participants, such as Participant O, expressed acceptance of robots but only if used properly according to them, saying: "I feel optimistic about robots. But not in my house. That means that they are doing things for the community…". There were no notable patterns in terms of the relationship between this theme and participants' reported experiences and attitudes.

Envisioning how robots can be used. This was one of the least prevalent themes (N = 5) that captured the specific ways in which participants envision that robots will be used in the future (e.g., "I definitely see potential for things like cleaners, like I'm sure, in 20 years' time …"). Although somewhat similar to the first theme, *Robots are helpful and useful*, this theme focuses specifically on participants sharing detailed accounts of ways in which they imagine robots will be used. For example, Participant N imagined the following scenario:

> I think having a little robot that can just do sort of simple tasks for you, which you may not want to do yourself, which obviously, it has the technology so it can just be quite helpful. If you just have a robot in your house, like in the same way that things like Alexa and like, all those little things that you can tell them, oh, remind me to do this or look up this thing for me, sort of thing…

Four of the participants shared similar scenarios immediately after being asked about the association which demonstrates a certain level of underlying acceptance of interaction with robots. This is supported by the fact that the sentiment the four participants expressed was entirely positive (see Table 3.14) and they all reported positive attitudes for the first subscale of the NARS, *Interaction with robots*. All but one of the participants made some reference to non-fictional experiences or knowledge, typically referring to existing robots that are not yet widely used or currently under development.

3.4.2.2.3 Emergent themes from the association of robots with "emotionless" and "cold"

Current state of robots. This was the second most prevalent theme (N = 8) that captured participants' views about robots currently being *emotionless* and/or *cold*. The sentiment expressed by participants for this theme was predominately neutral and demonstrated participants' knowledge about robots in a mostly factual tone (e.g., "The real life [robots], are still at the stage of being emotionless."). Both participants who stated that they would associate robots with being emotionless and/or cold and those who did not, contributed to the theme which was consistent with most participants' reported number of experiences with non-fictional robots. All eight participants also referred to experiences and / or knowledge relating to non-fictional robots (see Table 3.14) which further emphasises the factual nature of this theme. Interestingly, three participants pointed out that while robots are indeed emotionless, their direct experiences with robots did not necessarily allow them to think of robots as cold. For example, Participant C stated, "Emotionless, yeah. I wouldn't say cold. They can be easily programmed to be friendly and chirpy… ". Similarly, Participant E said that robots "… may not feel emotion, but they do have personality. So I wouldn't see them as cold.". Although this theme is not linked to participants' attitudes in an obvious way, it does indicate that most participants used their knowledge and / or experiences with real robots to inform their assessment of robots as emotionless and cold.

Robots not capable of having real emotions. This was the third most prevalent theme (N = 6) that encapsulated participants' beliefs that emotions were in some way reserved for humans (e.g., "… we have emotions. Ok, so that's really good. We have something that others or other mechanics or other things don't have…"); or that any emotion a robot may convey is artificial in some way (e.g., "… it's an artificial emotion that the robot is presenting, and that is very different from a real emotion."). The sentiment participants expressed was predominantly neutral (see Table 3.14) and only two participants, Participant H and L, were explicitly positive about robots not being capable of having emotions, stating their reasoning as the fact that they found the potential of emotional robots “scary” and felt that “… emotions are the most human thing.”. Their responses were closely linked to the last theme for this association, *Concerns and consequences* which better explains the relationship between this association and participants’ attitudes. Half of the participants also referred to knowledge about real robots, specifically in relation to robots being capable of displaying “artificial emotion”.

For example, Participant E shared that:

> They might not be able to understand or process well, they might be able to understand emotions, because I know there are robots that can pick up on your sort of facial expressions … You could tell a robot to feel sad when it meets certain parameters, but it's not the robot deciding that it sad itself. It's you've told it to.

Overall, there was no notable pattern between this theme and participants’ attitudes that could not be better explained by the next two themes. However, much like the first theme, Current state of robots, it does demonstrate that participants’ knowledge and /or experiences with non-fictional robots inform the association with *emotionless* and *cold*.

Envisioning a future with emotional robots. This was the most prevalent theme (N = 9) and it captured the various ways in which people imagined a future with robots that are capable of experiencing emotions. This was one of the most diverse themes in terms of the sentiment which participants expressed (see Table 3.14) which to a certain extent reflected participants’ reported attitudes for the second NARS subscale, *Social influence of robots*. It should also be noted that sentiment was not necessarily related to whether participants actually associated robots with the words *emotionless* and *cold*. Only two participant expressed entirely negative sentiment when envisioning robots with emotions and also had the top two most negative attitudes both for the second NARS subscale and overall. For example, Participant O expressed concerns about being able to control robots if they had emotions:

> I think of for me, I would prefer the robots to me emotionless. Not having human emotions at all just to do a specific technical function. Because if I imagine them being emotional, I just feel like it's crossing the line and crossing the boundary of how I can control it.

In contrast, Participant J who expressed more positive sentiment toward emotional robots and reported some of the most positive attitudes for all NARS subscales, was considerably less concerned about losing control:

> If … a robot that did have something like consciousness, there's no reason to think that it would just be like, be unable to display empathy and be completely cold, I think it would just be like humans, like it'd have a bigger array of emotions. … So I think that it could add value to a society.

The pattern of relationship between sentiment and attitudes was similar for the rest of the participants although most people acknowledged that emotion "could be a good thing in certain situations. … However, not every situation." and that whether they would like to see a robot with emotions "… would depend on the function of the robot". As such, the presence of this theme is not a definitive predictor of participants' attitudes but does indicate that particularly negative or positive outlook in relation to emotional robots impacts participants' attitudes toward the social influence of robots.

Concerns and consequences of having robots with emotions. This was the least prevalent theme (N = 5) and it captured the concerns that participants had in relation to emotional robots and / or their own preferences for emotionless robots. The sentiment participants expressed was on average almost entirely negative (see Table 3.14) unlike similar themes for the other associations.

The three participants with the most negative attitudes, both overall and for the second NARS subscales, contributed to this theme. Unsurprisingly, all three participants *strongly agreed* or *agreed* with the NARS statement, *Something bad might happen if robots developed into living beings*. However, the three participants did not necessarily agree with the statement, *I would feel uneasy if robots really had emotions*. The only obvious difference between those participants was their number of direct experiences that they reported, with the participant who had more experiences, also reporting less agreement with that particular NARS statement. In line with the previous themes, Participants H and L, who expressed that robots were not and should not be capable of emotion, also expressed concerns regarding such a possibility. For example, Participant L shared that, "… I don't know who I would trust then. If around me are robots and I cannot recognise that they are robots from their emotions.". The idea of robots being indistinguishable from humans was prevalent in this theme, "… if they're not emotionless, then what's the difference between (a] robot and a human?”. Fictional portrayals of robots may have a role in bringing up such concerns, as two of the participants referred to specific fictional examples (i.e., the film ‘Her’ and the game ‘Detroit: Become Human’) where robots are presented as convincingly human with arguably negative consequences. These participants also reported some of the highest number of recent experiences relating to fictional robots, although they also reported a high number of indirect (but not direct) experiences with real-life robots.

The other two participants, although still negative about the impact of emotional robots, spoke about practical concerns such as: “[If a robot] doesn't fancy getting to work because it feels a bit down. That's not really good for your work, is it, really?” and “If [a robot] feels a particular emotion, I mean, the robot is there for me to control it for a specific function of functions. So if it has emotions, I can't really control it.”. Interestingly, these two participants were very different in terms of the number of direct and indirect experiences with non-fictional robots that they reported and with their overall attitudes. The participant with more experience had positive attitudes and the participant with less experience had negative attitudes as measured by the NARS. However, they both still agreed with the NARS statement, *I would feel uneasy if robots really had emotions*.

Overall, this theme does demonstrate that the viewpoints participants expressed in the second theme are in some cases linked to real concerns regarding emotion in robots, which likely has some impact on people’s attitudes especially where the second NARS subscale is concerned. However, participants’ experiences with fictional and non-fictional robots may determine the extent to which participants’ concerns about emotional robots affect their overall attitudes.

3.4.2.2.4 Emergent themes from the association of robots with “artificial”

Defining the nature of robots. This was the most prevalent theme ($N = 9$) and it captured the way participants defined robots based on whether they believed the label *artificial* was applicable to robots or not. Participants spoke in two main ways about the artificial nature of robots, *robots as human creations* (e.g., “If we just consider artificial it means … someone has just created that.”); and *robots as real vs. not real* (e.g., “But robots themselves are not artificial. They are real. They are real robots.”). Nearly all participants expressed entirely neutral sentiment (see Table 3.14) when speaking about the artificiality of robots, with the exception of Participant G who strongly opposed to describing robots as artificial: “I think [artificial] kind of almost minimizes [robots] in some regards. … I kind of have a little bit of an issue with the word artificial. … It doesn't seem fit for purpose because they're very real…”

While this theme does not seem to be particularly informative in terms of participants’ attitudes, it is interesting that some people feel strongly about labelling robots as artificial (e.g., “No, no. They’re a very real thing.”) while others merely acknowledged that robots can be described as artificial (e.g., “It didn't immediately come into my brain, the word itself. But, yeah, I do understand that. And I do actually agree with it because it's kind of not organic…”). Combined with the fact that artificial was the least agreed with association, these findings point to a surprisingly diverse understanding of robots as artificial which was not obvious from Study 3.

Robots are artificial humans. This was the second most prevalent theme ($N = 6$) which again focused on how participants defined robots as artificial. However, unlike the previous theme, this one captured the idea of robots as artificial (or “fake”) humans (e.g., “… if I really associate it with artificial, it will be for this reason that they are like, fake versions, you know.”). Also unlike the previous theme, the influence of fiction was clearly present, with half of the participants referring to fictional portrayals of humanoid robots in some way and reporting multiple recent experiences with fictional robots. For example, Participant E stated, “… playing ‘Fallout’ [a game] with the synths, they are artificial humans with artificial intelligence. So, yeah, it's definitely an association I would make.”. The sentiment participants expressed in relation to this theme was predominantly neutral (see Table 3.14) and there were no notable patterns in the attitudes participants reported. However, it should be noted that the participants who expressed some negative sentiment toward the idea of robots as artificial humans both spoke about emotion (e.g., “I would say artificial in the way that it's not, because unlike a human, it doesn't have, like, the capacity to, like, empathize or anything. Like, they're not sentient.”). This may point to a link between the association with *artificial* and the association with *emotionless* (see Section 3.1.2.2.4).

Concerns and consequences regarding technology advancement and use. This was the third most prevalent theme ($N = 5$) that captured the concerns participants had both in relation to mislabelling robots as artificial (e.g., “… they're very real and if they're used wrong, I feel it could be quite a negative, catastrophic result from it.”) and the consequences of using ‘artificial’ robots (e.g., “… if you just introduce these things, these devices or these artificial robots, [it would] just make it really harder for the people.”). The sentiment participants expressed in relation to robots was predominantly negative (see Table 3.14) and two of the participants referred to fiction in their responses, following on from the previous theme of robots as artificial humans. Interestingly those were also the only two participants who did not express entirely negative views toward the consequences of using robots. For example, while Participant E said that “if things happen the same way they did in sci-fi, then yeah, it would be a very terrifying thing if you were walking around not knowing [if] the person next to you [was] … fake”. However, Participant E also stated that the association with robots being artificial humans did not change how they felt about robots “… particularly because I know it's just sci-fi. If they were to be real, then I would be terrified. But then I'm quite rational and just know that, yeah. We're nowhere near that.”. Unsurprisingly, both participants who expressed similar opinions reported two of the highest numbers of experiences with both fictional and non-fictional robots. Similar to the second theme for the association with *artificial intelligence*, these findings once again suggest that fictional portrayals of robots are influential in raising concerns about robots but may not necessarily result in negative attitudes if the individual has knowledge and experience with real robots.

Artificiality of robots is a good thing. This was the least prevalent theme (N = 4) that encapsulated participants’ preference for the artificiality of robots in terms of how they defined the term. For example, Participant L stated that the association was “very positive” for them “because, for me artificial is again is related to, to being non-emotional. So, again if a robot with emotions, that will be a hard thing to deal with. So I think they are artificial.”. Participant N similarly shared that associating robots with artificial was “a good thing, because I think if robots were, you know, sentient, you know. Or like if they could think and feel, I think people would … get too attached to them.”. Unsurprisingly, when compared to the other two participants in this theme who were mostly neutral about their preference for artificial robots, Participant L and N had more negative attitudes overall and more specifically, had notably negative attitudes for the second NARS subscale, *Social influence of robots*. Similar to the second theme, these findings may point to a link between associating robots with *artificial* and with *emotionless* (see Section 3.1.2.2.4).

3.4.2.2.5 Emergent themes from the association of robots with "machines"

Defining the nature of robots. This was the second most prevalent theme across all associations (N = 12) that consisted of mostly participants who stated that they would associate robots with the word *machines*. This theme encapsulated the different ways in which participants utilised the meaning of the word machine in order to define robots and identify similarities between the two. Two main sub-themes were also found: half of the participants (N = 6) were *contradictory or unsure about the definition of a robot* (e.g., "I'm not really sure what the definition is of a machine.") and some participants (N = 5) explicitly agreed that *robots are machines* (e.g., "I think every machine is a robot as well. Because it follows your command."). The majority of participants (N = 8) referred to their knowledge of and / or experiences with real robots during this interview segment and this was true for participants with varying numbers of experiences relating to both fictional and non-fictional robots. Interestingly, four participants used their knowledge to highlight that whether a robot is defined as a machine or not may be dependent on its appearance. As Participant K explains:

> … I think people would call a humanoid robot a robot more than a machine. Although I suppose a human a humanoid robot is still a machine, isn't it? Because it looks like a humanoid you probably wouldn't immediately associate that with the machine. Whereas the production line robots; you would think that they were machines.

Similarly, Participant N pointed out that a robot may be perceived differently depending on what people imagine a robot to be: "… if you think that a robot is like something metallic …then you might think, OK, that's a machine. If you think of robot that's sleek and trendy, then you might think, OK, it's completely different from a machine."

Appearance was not the only characteristic that participants were conflicted about in terms of whether robots were machines. For example, the function and autonomy of robots was also subject to evaluation. For example, Participant A was unsure about what could be considered a robot:

> … machines are things like the washing machine and the photocopier, they haven't got arms. But they are also programmable, so. What is [it]? Because, you know, you load your washing machine and program it to do the job. You know, is it a robot because

it's robotically doing a job for you? It does the same job, it's programmed to do the same thing.

Overall, it is unclear how and if this theme relates to participants' attitudes especially given that participants predominantly expressed neutral sentiment (see Table 3.14) and approached this topic with a mostly descriptive and questioning tone. However, it is clear that although most participants do associate robots with machines, the reasons for which they do so are varied and not necessarily explained by their experiences with real robots.

Robots distinct from machines. This was the second most prevalent theme (N = 7) that captured the way participants differentiated robots from machines. All three participants who stated that they would not associate robots with machines contributed to this theme. Robots were distinguished from machines based on their *appearance* (e.g., "They aren't humanoid or anything. They are machines."), *intelligence* (e.g., ". I think if we used the word machine or robot, I think of the robot being something more intelligent, I guess"), *autonomy* (e.g., "… robots have more autonomy than machines"), and *function* (e.g., "Robots with a singular purpose seem to become more machine than robot in my mind, I think."). There was considerable co-occurrence between this theme and the previous one which is not surprising given that most participants were conflicted about whether robots and machines could be considered the same. Similar to the previous theme, all but one participant referred to knowledge and / or experience with real robots to identify differences between robots and machines. However, there was no apparent relationship with the number of experiences participants reported, indicating that the way participants differentiate robots from machines is not necessarily determined by their experiences. Participants expressed predominantly neutral sentiment (see Table 3.14) and used largely descriptive language when talking about the differences between machines and robots. Additionally, there was no notable difference in participants' attitudes between participants who associated robots with machines and those who did not.

Acceptance of robots as machines differs. This was the third most prevalent theme (N = 6) which the extent to which associating robots with *machines* was a positive or negative according to participants. This was one of the most diverse themes in terms of the sentiment which participants expressed, although still predominantly negative (see Table 3.14). Half of the participants felt that viewing robots as machines was a negative for them. For example, Participant F stated that for them "… [machines] is not a positive term. Because it kind of makes you think a machine is something that doesn't … have a consciousness. It doesn't have a soul. And that can be quite scary …". Other participants were more neutral about their responses, such as Participant D who shared that "… robots occupy a slightly different category [than machines]. … Like something more fun or more intelligent …". Only Participant O felt strongly that associating robots with machines was a positive thing for them as it would allow them to control robots as they would control other machines; "… being able to see and understand and like kind of have control over [robots] rather than …just doing things that humans can't really control". Despite the variety of sentiments and reasons participants expressed, there were no notable differences that would indicate this theme was influential in shaping participants' attitudes.

3.4.3 Discussion

3.4.3.1 The link between artificial intelligence, computing, and attitudes toward robots

Participants' association with artificial intelligence and computing represented Module 1, or the core aspects of the social representation of robots. This was the only association that all participants in Study 3 agreed that they would make and as such was consistent with the findings from Study 3 as it was expected that associations relating to the largest module would be more common in the population. The thematic analysis of interview segments relating to this association supported the overall conclusions drawn from Study 3 (see Section 3.3.3.4) as it demonstrated the role that AI and computing play in people's understanding and representation of robots. Three themes emerged from the data and, on average, participants referred to their experiences with, and knowledge of, non-fictional robots the most for this association. This finding supports the idea that Module 1, and a substantial part of the core of the representation of robots, is a product of the historic ties between the field of artificial intelligence and robotics. It also indicates that there is some common or shared knowledge among individuals that enables them to identify AI and computing as an integral part of robotics.

On average, the sentiment participants expressed for this association was neutral which is in line with the overall neutrality of the associations that made up Module 1 in the semantic network (see Section 3.3.2.1). Not all participants viewed the link between AI and robots in the same way; as demonstrated by the emergence of two themes that encapsulated the consequences and the benefits of AI that participants could foresee. The link between participants' attitudes and whether they spoke about the consequences or benefits of AI (or both) was not straightforward and it appeared that participants' experiences played a large role in the relationship. There was evidence that participants' concerns were at least partially driven by their experiences with negative fictional portrayals of robots or AI. Having said that, it appeared that said negative portrayals may not have affected individuals' attitudes as participants recognised that such fictional representations were not consistent with their own knowledge about real robots and AI. However, for participants who did not have extensive knowledge and experience with non-fictional robots, the impact of fictional portrayals on their concerns (and subsequently attitudes) may have been greater. In contrast, the theme that captured participants' perception of how AI can improve the usefulness of robots appeared to be primarily driven by participants' experience with non-fictional robots. The link between fiction and the association with artificial intelligence is supported by the finding that Module 1 was where almost all associations relating to fiction were located in the semantic network. According to these findings, associating robots with AI and computing is unlikely to play a large role in shaping people attitudes directly. However, participants' responses to the NARS statements relating to artificial intelligence and sentience could be susceptible to concerns that are not contradicted by individuals' knowledge about robotics.

3.4.3.2 The link between the usefulness and attitudes toward robots

Participants' association with the words useful and helpful represented Module 2, or the potential of robots. Most, but not all, participants agreed that they would make the association which was consistent with the findings from Study 3 as it was expected that associations relating to the second largest module would be prevalent in the population. The thematic analysis of interview segments relating to this association supported the overall conclusions drawn from Study 3 (see Section 3.3.3.4) as almost all participants acknowledged that robots are, or could potentially be, useful and helpful. This was true regardless of participants' overall attitudes as demonstrated by the first of four themes that emerged from the data. Many participants also expressed concerns about the usefulness of robots, as indicated by the second theme. These concerns were primarily cantered around job loss for humans and the consequences of depending on robots in the future. Unlike the concerns surrounding artificial intelligence, the consequences participants spoke about in regards to this association were almost exclusively supported by references to participants' experience and knowledge with real robots. Automation in the context of manufacturing was primarily evoked and may indicate that manufacturing is a particularly vivid example of real robots and their usefulness (e.g., being more efficient and precise) as well as their negative impact. This was consistent with the findings from Study 2 and 3 as manufacturing robots were some of the most salient examples of non-fictional robots in both studies.

In terms of attitudes, participants' concerns in relation to the usefulness of robots may have been particularly influential for NARS statements relating to interaction with robots. For many participants, the current and potential usefulness of robots was the primary focus of the association and their concerns were embedded in an overall positive narrative about robots as useful for humans. Those participants had more positive attitudes overall and were also less negative toward working with robots than participants whose concerns appeared to be more salient. This difference could be explained by the fact that the participants who had more positive attitudes also reported more direct experiences with robots, which has previously been shown to correlate with more positive attitudes (Bartneck et al., 2006). Similar to the previous theme, it appeared that being aware of the potential consequences of using robots (such as job loss) was not necessarily a sign that participants had negative attitudes. Overall, associating robots with being useful and helpful yielded almost entirely positive sentiment toward robots regardless of whether participants expressed any concerns. These findings were in line with the results of Study 3 (see Section 3.3.3.4) as Module 2 was found to contain predominantly positive associations (e.g., "efficient", "clever"); associations relating to manufacturing (e.g., "automation"); and two negative associations (i.e., "scary" and "non-human). For some participants, the extent to which robots could be defined as helpful to humans was conditional on the responsible use of robots as evidenced by the third theme that emerged from the data. This theme was linked to the concerns some participants expressed but also highlighted a general worry about how useful technology is currently being used. This was a particularly interesting finding as one of the larges nodes for Module 2 was "technology", a relatively neutral association that was strongly linked to associations such as "helpful", "advanced", and "efficient".

The most influential theme in terms of participants' attitudes was also the least prevalent one for this association. It captured the ability of participants to not only recognise the usefulness of robots but also be able to envision, unprompted and in detail, how robots would be used in the future. All but one of the participants whose response fell under the last theme imagined a positive scenario in which robots could be useful. This was linked to very positive attitudes toward interaction with robots (i.e., second NARS subscale), perhaps indicating an underlying acceptance of the use of robots. Alternatively, being able to envision negative scenarios in addition to expressing concerns may be detrimental to people's attitudes toward interaction with robots. However, the data in Study 4 did not provide any further insights on this topic.

3.4.3.3 The link between emotionless robots and attitudes toward robots

Participants' association with the words emotionless and cold represented Module 3, or the idea of robots as mechanical and unfeeling constructs. Most, but not all, participants agreed that they would make the association which was consistent with the findings from Study 3 as it was expected that associations relating to the third largest module would be prevalent in the population but less so than for the first two modules. The first theme that emerged from the analysis captured participants' knowledge about real robots and their current state in terms of emotion. Participants' statements were predominantly descriptive, acknowledging robots as lacking emotion in a manner that was not directly linked to attitudes. Some participants acknowledged that robots were emotionless but not necessarily cold as they could be programmed to appear friendly. This finding may explain why in Study 3, "emotionless" was a modular hub (i.e., a dominant association) while "cold" was not. Furthermore, every participant who spoke about the current state of robots referred to their experience with and / or knowledge of non-fictional robots. This finding suggests that some participants initially relied on such experiences to make a judgement about the extent to which robots are emotionless and cold. The second theme further demonstrated that non-fictional experiences played a role in the association but with the added sentiment that robots cannot have real emotions, or at least have emotions in the same way that humans do. Although most participants expressed predominantly neutral sentiment, some identified robots as a potential threat to the qualities (such as emotion) that participants thought to be inherently reserved for humans. It was not clear how and if such perceived threat was related to participants' attitudes although it may explain some participants' aversion to robots potentially being capable of emotion.

Most participants were able to envision a future with emotional robots regardless of whether they made the association with emotionless or not. These imagined scenarios made up the third theme that was by far the most varied one in terms of sentiment across all the associations. Not only was this the most prevalent theme but it also captured the diversity of participants' attitudes toward the social influence of robots (i.e., second NARS subscale). While some participants envisioned how robots with emotion could be useful and ultimately desirable, others rejected the idea for various reasons. The results suggested that participants who are more open to encountering emotion in robots (be it genuine or programmed) also have overall more positive attitudes. This finding was in line with the results of Study 3 which suggested that Module 3 represented the conflicting qualities of robots that likely reflected the variety of fictional and non-fictional robots that people were exposed to. However, participants did not explicitly refer to their experiences with fictional and non-fictional robots although that may not necessarily mean that these experiences did not influence participants.

What participants envisioned was linked to the final theme that captured their concerns regarding emotional robots. This was especially true for the participants that held the most negative attitudes overall and for the second NARS subscale in particular. Generally, participants were concerned about robots becoming indistinguishable from humans, a topic that is often portrayed in fiction in an arguably negative way. There was some evidence that fiction indeed played a role in bringing up such concerns, although not all participants directly referred to their reported recent experiences with fictional robots. Similar to the concerns participants expressed for the other associations, direct experiences likely played a role in counteracting fictional depictions. This was especially true for NARS statements relating to participants' feelings during interaction. The participant with most direct experiences but little recent fictional experiences rejected the association with emotionless, had the most positive attitudes, and overall perceived robots as friendly and engaging. These findings support previous research that indicated a positive correlation between direct interaction with robots and attitudes (Bartneck et al., 2006).

"Emotionless" appeared to be the association with the most direct link to participants' reported attitudes, especially for the second NARS subscale. It was also the most diverse association in in terms of the sentiment participant expressed across the four themes. This was consistent with the results of Study 3 as Module 3 contained a mixture of negative, positive, and neutral associations. In Study 3 “emotionless” was an entirely negative association. However, the interviews for Study 4 suggest that the interpretation of this valence is not straightforward. For some participants, the association was clearly positive in the sense that they preferred robots to not express or have emotion. Others saw emotionless as a negative association and preferred to see emotional robots. Overall, this association captured something about how participants envisioned robots in the future to be. More specifically, individuals who were more accepting of the possibility of robots as independent social actors generally had more positive attitudes, especially attitudes toward the social impact of robots.

3.4.3.4 The link between labelling robots as artificial and attitudes toward robots

Participants’ association with the word artificial represented Module 4, or the idea of robots as artificial constructs. Less than half of the participants agreed that they would make the association which was consistent with the findings from Study 3 as it was expected that associations relating to the fourth largest module would not be as prevalent in the population. In fact, this was the association that was met with the least explicit agreement although even participants who did not necessarily state that they would make the association understood why artificial was a label that could be applied to robots. People defined robots in terms of their artificiality in two main ways as captured by the first theme: (1) as object made by humans, and (2) as “real” vs. “not real”. The first interpretation of artificial is somewhat in line with the findings from Study 3 as Module 4 contained associations with other human made objects (e.g., “tool”, “toy”). The second definition was somewhat more ambiguous but participants who did not agree with the association generally considered artificial to be an antonym for “real”. The second theme was similarly about understanding why someone may consider robots to be artificial; namely, robots were defined as artificial (or “fake”) humans which may explain the presence of the association “humanoid” in Module 4. Unlike the first theme, this theme was likely influenced by fictional portrayals of humanoid robots as participants primarily referenced their experiences with fiction. There was also a link between artificial and the previous association, emotionless, as some participants considered robots to be artificial only if they possessed emotions. Further support for the link between artificial and emotionless was provided from the forth theme where artificiality was framed as a desirable property of robots cantered around the idea of non-emotional (i.e., artificial) robots being preferable. As for emotionless, participants whose responses fell under this theme had negative attitudes toward the social influence of robots (i.e., second NARS subscale). The third theme that captured participants concerns in regards to robots as artificial was similar to the concerns participants had for AI in that they were predominantly driven by experiences with fictional robots but counteracted by participants' knowledge about real robots. This finding supports the idea that acknowledging concerns related to robots is not necessarily linked to negative attitudes if participants have sufficient knowledge about the real impact of robotics. As such, associating or not associating robots with being artificial is not particularly useful for predicting participants' attitudes although artificial may be closely linked to the association with emotionless.

3.4.3.5 The link between machines and attitudes toward robots

Participants' association with machines represented Module 5, or the idea that robots possess characteristics typically associated with machines. Most participants agreed that they would make the association which was surprising as it was expected that associations relating to the smallest module would not be as prevalent in the population. However, the first theme that emerged from the data demonstrated not only the variety of ways in which people saw robots as similar to machines (e.g., both perform a specific function) but also participants' contradictory opinions on the definition of robots. The second theme supported the contradictory definition of robots as it highlighted the ways in which participants thought robots were different from machines and the ways in which robots may appear less or more machine-like (e.g., humanoid appearance and more autonomy). These themes support the idea that robots are likely perceived as a distinct social category unlike conventional machines. The third and final theme suggested that the extent to which participants accepted robots when the label machine was applied to them differed. This theme was linked to some participants' preference for non-emotional robots (i.e., machines) and other participants' preference for more humanised and social robots. The findings reflect the structure of the semantic network as Module 5 was found to contain contradictory in valence associations such as "powerful" and "uncanny" cantered around the neutral association "machine". The sentiment participants expressed across the three themes was also mixed although predominantly neutral. There was no evidence that associating robots with machines (or not) was directly linked to participants' attitudes as measured by the NARS. However, "machines" may be an association that indicates how people are likely to view robots as a category (e.g., machine-like and emotionless or human-like and social).

3.4.3.6 Limitations

A combination of traditional thematic analysis and detecting patterns between participant's sentiment and reported NARS scores was used to analyse the data. In order to simplify the findings of the qualitative analysis, *t*-tests were conducted to compare the attitudes of participants who endorsed or did not endorse the associations. Due to the fact that nearly all participants endorsed the association which represented the first two modules, this was only possible for the associations relating to Modules 3, 4, and 5. It was assumed that endorsing or not endorsing a particular association equates to that particular concept (i.e., module) being part of (or not) in an individual's representation of robots. Given the obvious limitations in terms of sample size, these findings should be treated as preliminary support for the relationship between representations and attitudes.

Furthermore, some attention should be given to the logic behind selecting the associations representing each module and what the implications of that choice could be. The associations selected were one or two of the largest associations for a particular module and were typically identified as part of the core of the social representation. It was thought that they would best represent each module as they are highly connected with all or nearly all other associations in that particular module and thus maintain the modular structure of the representation. Although this makes sense based on Abric's work no literature informed this selection method. In regards to the methodology of Keczer et al.'s semantic network on which Study 3's methodology was largely based, that it was prior qualitative work on the concept that they were studying that drove the aims of their study, as well as the interpretation of their modules and overall findings. In that respect, the work presented in Chapter 3 is primarily explorative in nature which is not necessarily a problem given that there is very little work done on the social representation of robots. However, caution should be applied when interpreting the findings from those studies.

3.4.3.7 Conclusion

Study 4 presented initial evidence for the relationship between the way people represent the concept of robots and their explicit attitudes. However, this relationship does not appear to be straightforward and people's experiences with both fictional and non-fictional robots likely play a significant role in how the representation of robots is shaped and influences people’s attitudes. Although Study 4 provided insight into the impact of fictional depictions of robots on people’s representations, it is unclear how impactful those depictions are in terms of people’s attitudes especially when compared to experiences with real robots. As such, Chapter 4 and 5 present a series of experimental studies that aimed to investigate whether the perceived fictionality of robots (fictional vs. real) affected people’s attitudes toward specific robots and robots in general.

Chapter 4: Manipulating the Fictionality of Robots

Chapter Summary

Chapter 4 presents three pilot studies that had the primary aim of developing a practical method of manipulating the perceived fictionality (fictional vs. non-fictional) of identical robots for use in Study 5. Each pilot study tested videos depicting social robots and complimentary paratext which was utilized as a means of manipulating participants' belief in whether the video and robots were fictional or not. Chapter 4 also briefly introduces the concepts of *fictionality* and *perceived realness* which are relevant to the experimental work presented in Chapter 5.

4.1 Introduction

As established in Chapter 3, fiction has a place in both the social and individual representation of robots. However, the direct impact of fictional depictions of robots has yet to be directly investigated especially when compared to the impact of indirect exposure (e.g., seeing a robot on the news) to real robots. This chapter presents three pilot studies that aimed to inform the materials and design used in Study 5 (see Section 5.2). The pilot studies explored how labelling media as fact or fiction (in this case videos depicting robots) could influence individuals' perception of the fictional status of both the robots and videos.

Literature of the impact of fiction on attitudes primarily focuses on *framing*. In other words, how certain social groups, issues, or events are presented in fiction (typically in a positive or negative way). Direct test of fictional framing shows that fiction can influence people's attitudes and beliefs on a variety of different issues (e.g., van den Bulck, 2002; Green et al., 2004; Mulligan & Habel, 2011; Johnson et al., 2013). From existing research (see Section 1.5) we know that the way robots are framed in fiction (as "good" vs. "bad") does not necessarily lead to positive or negative attitudes. Instead, it leads to confused, ambiguous, double-minded attitudes that are not necessarily driven by real information as most people do not encounter real robots directly or indirectly in non-fictional media (Bruckenberger et al., 2013). As such, the pilot studies presented in this chapter do not use *framing* or attempt to actively persuade participants that robots are "good" or "bad". Instead, an attempt is made to manipulate the *fictionality* of the robots and media they appear in. Fictionality is defined not as an inherent attribute of fiction or non-fiction but rather as the act of intentionally communicating invention via direct (e.g., labelling a book as fictional) and indirect (e.g., telling a story about flying humans) contextual cues (Mendelson & Papacharissi, 2007; Busselle & Bilandzic, 2008).

As such, the primary aim of this chapter was to develop a method of manipulating participants' belief in the fictionality of robots that they were unlikely to be familiar with. This was done by intentionally signalling the video of the robots as fictional or non-fictional via textual information provided prior to the video that stated whether the scenario and robots were "invented" or real.

4.2 Pilot Study 1 - Manipulating the Perceived Realism of Robots

Pilot Study 1 investigated whether it was possible to manipulate what participants believed about the robots depicted in two videos - one depicting a fictional robot, the other depicting a non-fictional robot – by providing information implying that the video and/or robots were fictional or non-fictional (or no information was provided).

The two videos were selected as it was thought that they would provide a good range of depictions of robots with which we could test the selected method and were not compared to each other. It was expected that there would be some difference between the conditions in participants' ratings of the perceived realism of the video and robots. Video quality was also measured to inform the selection of the video for the main study but no difference between the conditions was expected.

4.2.1 Method

4.2.1.1 Participants

A sample of 41 Psychology students at the University of Sheffield completed the pilot study.[9] First year undergraduates (N = 25) were recruited via the Department of Psychology's Online Research Participation System (ORPS) and rewarded with course credits as compensation for their time. Postgraduate students (N = 16) were recruited using poster displayed at the Department of Psychology and received no reward for completing the study. Data from participants who indicated that they had seen the videos or/and robots before taking part in the study were excluded from further analysis (see Table 4.1).

[9] A priori power analysis was not conducted for this study due to an oversight by the author.

Table 4.1

Total Number of Valid and Excluded Cases for Each Video

	Video 1	Video 2
Valid	37 (90.24%)	36 (87.80%)
Excluded	4 (9.76%)	5 (12.20%)

Note. Video 1 refers to the ‘BUDDY: Your Family’s Companion Robot’ video; Video 2 refers to the ‘Robot & Frank’ video.

4.2.1.2 Procedure and Materials

Data was collected online via Qualtrics. Participants were told that the aim of the study was to investigate the way that advanced technology (e.g., robots) is portrayed in the media (e.g., in movies). For each video, participants were randomly allocated to one of three conditions and received different descriptions of the robots that they would see in the video (see Table 4.2). For example, a participant could have seen the first video and been told that it is fictional and then received information for the second video implying that it is non-fictional. Participants in the non-fictional condition received information implying that the robots in the two videos were real. While participants in the fictional condition received information implying that the robots in the two videos were fictional and participants in the control condition received information that did not specify whether the robots were real or not.

Table 4.2

Key Information Given to Participants in the Fictional, Non-Fictional, and Control Conditions

Condition	Information
Fictional	**Video 1: 'BUDDY: Your Family's Companion Robot'** The video you are about to watch is a short fictional film created as part of a student project. The video depicts the interaction between a fictional robot called BUDDY and a human family. For the purposes of the video, a remote controlled toy was used to portray BUDDY and its interaction with professional actors. **Video 2: 'Robot & Frank'** The video you are about to watch is from the movie 'Robot & Frank' (2012). The video depicts the interaction between the two main characters, ROBOT and Frank. For the purposes of the movie, the character of ROBOT is played by an experienced human actor in a robot costume.
Non-fictional	**Video 1: 'BUDDY: Your Family's Companion Robot'** The video you are about to watch is a promotional video for an upcoming real-life robot. The video depicts the interaction between BUDDY and its potential users. For the purposes of the video, BUDDY was filmed interacting with a real family in their home. **Video 2: 'Robot & Frank'** The video you are about to watch is from the movie 'Robot & Frank'. The video depicts the interaction between the two main characters, ROBOT and Frank. For the purposes of the movie, the character of ROBOT is played by a real robot developed by one of the leading robotics companies.
Control	**Video 1: 'BUDDY: Your Family's Companion Robot'** The video you are about to watch was obtained from a publicly available source. The video depicts a robot called BUDDY. **Video 2: 'Robot & Frank'** The video you are about to watch was obtained from a publicly available source. The video depicts a robot called ROBOT.

Videos. Following the information, participants were asked to view the two videos in a random order. One of the videos was an official promotional video for a companion robot called 'Buddy' obtained directly from YouTube. The second video was created using Adobe Premiere Pro by selecting scenes from the movie 'Robot & Frank' (Schreier, 2012) and merging them together. The videos were edited so that they were both approximately 3-4 minutes long. All participants saw the same two videos in a random order regardless of the condition to which they were assigned.

Additional questions. Immediately after each video, participants were asked to indicate whether they had seen the robot and / or the video prior to taking part in the study.

Perceived realism of the video and robots. In order to assess the quality of the video, and the extent to which participants found the video and the robot to reflect reality, a self-report measure was developed specifically for Pilot Study 1 (see Table 4.3). The measure consisted of 10 items and each statement was rated on a 5-point scale (1 - *strongly disagree*, 2 – *disagree*, 3 – *undecided*, 4 – *agree*, 5 – *strongly agree*). This measure consisted of three subscales and the final score was the average of all items in each subscale. The order in which the statements appeared was randomised and the scale appeared at the end of the study before participants were debriefed. Participants were then debriefed and given the chance to share any additional comments they had in regards to the study (see Appendix D for debrief procedure). The implications of participants' feedback for the study's methodology were considered and discussed in the next sections.

Table 4.3

Subscales, Items, and Cronbach's α of the Perceived Realism Measure

Subscale	Item	Cronbach's α	
		Video 1	Video 2
Robot Realism	[Robot name] seems to be representative of current technology.	.63	.65
	I found [robot name]'s abilities believable.		
	[Robot name] was completely unrealistic. [a]		
Video Quality	I think that the video was well made.	.86	.72*
	The quality of the video was poor. [a c]		
	The video looked professionally shot.		
Video Realism	I think that the video was unrealistic.	-	-
	The video looked like it was part of a movie. [b c]		
	The video looked like it was part of a documentary. [b c]		
	The video looked like it was part of an advertisement. [b c]		

[a] Denotes reverse-scored questions.

[b] Denotes items that were removed in order to improve the overall reliability of the scale for Video 1.

[c] Denotes items that were removed in order to improve the overall reliability of the scale for Video 2.

* Cronbach's α after item deletion.

4.2.2 Results

The results are reported separately for each video and were not compared as this was not part of the aim of Pilot Study 1. A one-way ANOVA with three planned contrasts (fiction vs. non-fiction, fiction vs. control, non-fiction vs. control) was conducted for each video and for each dependent variable (robot realism, video quality, video realism). Assumptions of homogeneity of variance and normal distribution were also checked for each ANOVA (see Appendix J for details).

4.2.2.1 Video 1: 'BUDDY: Your Family's Companion Robot'

There was no significant effect of manipulating fictionality on how realistic participants thought the robot was, $F(2, 34) = 2.19$, $p = .127$, $\eta^2_p = .06$, 90% CI [0.00, 0.26]. There was also no significant effect of manipulating fictionality on how realistic participants thought the video was, $F(2, 34) = 0.83$, $p = .443$, $\eta^2_p = .02$, 90% CI [0.00, 0.16]. The way participants perceived the quality of the video was also not significantly affected by manipulating the fictionality of the video, $F(2, 34) = 1.43$, $p = .253$, $\eta^2_p = .04$, 90% CI [0.00, 0.21]. See Table 4.4 for descriptive statistics.

Additional feedback from participants. Six participants left additional comments in response to watching the *Buddy video*. Four participants mentioned the fact that the *Buddy video* is dubbed and two mentioned viewing adverts prior to the video. Two participants questioned the usefulness of Buddy and compared it to a smartphone that can move.

4.2.2.2 Video 2: 'Robot & Frank'

There was no significant effect of manipulating fictionality on how realistic participants thought the robot was, $F(2, 33) = 0.16$, $p = .856$, $\eta^2_p = .004$, 90% CI [0.00, 0.06]. There was also no significant effect of manipulating fictionality on how realistic participants thought the video was, $F(2, 33) = 2.29$, $p = .117$, $\eta^2_p = .06$, 90% CI [0.00, 0.27]. The way participants perceived the quality of the video was also not significantly affected by manipulating the fictionality of the video, $F(2, 33) = 0.26$, $p = .771$, $\eta^2_p = .008$, 90% CI [0.00, 0.09]. See Table 4.4 for descriptive statistics.

Additional feedback from participants. Two participants left additional comments in response to watching the Robot video. One participant commented that they found the character of ROBOT to be obviously a human in a costume and doubted whether any information would be sufficient to convince someone otherwise. The other participants commented that while their own lack of knowledge about robotics made them skeptical about the realism of the video, they could imagine a similar scenario in the future and as such thought the video had realistic elements.

Table 4.4

Mean, Standard Deviation, and 95% Confidence Intervals for Each Video, Condition, and Measure

Measure	Condition	*M* [95% CI]	*SD*
Video 1: 'BUDDY: Your Family's Companion Robot'			
Robot Realism	Fictional	3.25 [2.67, 3.83]	0.92
	Non-fictional	3.77 [3.48, 4.06]	0.48
	Control	3.83 [3.32, 4.35]	0.81
Video Realism	Fictional	2.58 [2.16, 3.00]	0.66
	Non-fictional	2.87 [2.65, 3.08]	0.36
	Control	2.65 [2.22, 3.07]	0.67
Video Quality	Fictional	2.96 [2.20, 3.72]	1.20
	Non-fictional	3.46 [2.75, 4.18]	1.18
	Control	3.71 [3.13, 4.29]	0.92
Video 2: 'Robot & Frank'			
Robot Realism	Fictional	3.08 [2.61, 3.55]	0.74
	Non-fictional	3.00 [2.47, 3.53]	0.24
	Control	2.89 [2.27, 3.50]	0.97
Video Realism	Fictional	2.17 [1.71, 2.62]	0.72
	Non-fictional	2.92 [2.34, 3.49]	0.90
	Control	2.67 [2.04, 3.29]	0.28
Video Quality	Fictional	4.00 [3.44, 4.56]	0.88
	Non-fictional	4.21 [3.77, 4.65]	0.69
	Control	4.08 [3.76, 4.41]	0.51

4.2.3 Discussion

The findings of Pilot Study 1 indicated that participants' ratings of the perceived realism and quality of the videos and robots were not significantly affected by whether the robots and videos were presented as fictional or non-fictional. As such, the methodology used was not sufficient to manipulate the perceived fictionality, or more accurately, the perceived realism, of the robots using the method and materials described in this pilot study. However, Pilot Study 1 had a number of limitations with respect to the instructions, videos, and measures which have been discussed below.

4.2.3.1 Limitations of the condition-specific information

Although it was initially thought that simply participating in a research study should provide sufficient context to satisfy the pragmatic nature of communicating fictionality, this idea may have been too simplistic. The information designed to manipulate the perceived fictionality of the video and robots was not necessarily suitable for the videos that were presented. This was indicated by a disagreement between the comments some participants left in regards to whether the manipulation was too obvious or not obvious enough. One participant in particular pointed out that the character in the 'Robot & Frank' video was obviously an actor in a suit and as such the information they were provided with would never convince them otherwise. Additionally, another participant thought that more emphasis should be put on whether the robots were fictional or not and pointed out that they did not pay much attention to the information that was provided. It is likely that while the methodology used in Pilot Study 1 was enough to 'intentionally signal' fictionality, it failed to account for other contextual cues such as elements in the videos that indicated the fictional status of the video and robots. In addition, it may be that the design of the study was in itself not efficient in emphasising and drawing attention to the condition-specific information.

4.2.3.2 Limitations of the video

In addition to the above mentioned limitations, the videos used in Pilot Study 1 also have a number of issues. With regards to the 'Buddy' video, participants mentioned the dubbing as distracting. Dubbing, especially when perceived as low in quality, could contribute to the overall quality evaluation of the video and, more importantly, affect realism evaluations as a consequence (Wissmath et al., 2009). Furthermore, due to the original promotional purpose of the video, there was a narration throughout, explaining the various functions and purposes of the Buddy robot. While it is not possible to know what participants thought about the style of the video, it could be that the narration rendered the video more advert-like than necessarily appropriate for the purposes of Pilot Study 1. This is to some extent supported by the consistently high scores participants gave the video when asked how similar to an advert it was. The second video, depicting the movie character Robot, may also have been problematic. Specifically, despite selecting movie scenes with limited dialogue, the video contained multiple references to other characters, changes in location, and changes in the behaviour of Robot and Frank, all of which were indicative of an underlying narrative which is a common feature in both fictional and non-fictional media. The process of scene selection and editing resulted in the breaking-up of said narrative and a potential loss of coherence where character development and storyline were concerned. Considering that in both the fictional and non-fictional condition, the video was presented as a work of fiction (i.e., a movie) and that the coherence of fictional narratives is a predictor of perceived realism (Cho et al., 2014), the effect (if any) of disrupting that narrative remains unclear. Additionally, there was some indication by participants comments that Robot was perceived as too unrealistic to be considered non-fictional; although, this could be dependent on participants' previous experience with fictional and non-fictional representations of robots.

4.2.3.3 Limitations of the measures

A self-report measure of perceived realism was developed for the purposes of this study. However, the unvalidated measures (robot realism, video quality, and video realism) and the overall poor reliability of the scales means that, if available, existing measures would likely be more suitable for detecting whether the manipulation was successful. In addition, participants were not explicitly asked if they guessed the purpose of the study or if they were aware that they were being deceived as to the nature of the videos. Given the additional comments some participants left, this pilot study would have benefited from a formal check of whether participants were naïve to the manipulation.

4.2.3.4 Changes to the self-report measures of perceived realism

Some studies have used a single Likert scale item to assess participants' perception of the overall realism video-recorded HRI scenarios (e.g., Woods et al., 2006) that are similar to one of the original items in Pilot Study 1, "I think that the video was unrealistic". However, it seems unlikely that realism is a unidimensional construct that can be captured by a single item. For example, Green (2004) makes a distinction between the subjective evaluations individuals make in regards to the plausibility of a story (i.e., are the characters, events, and settings in the story are *like* those in real-life) and the more objective truth value of the story (i.e., the characters and settings in the story are *factual representations* of present or past events). Cho et al. (2014) have further split perceived realism into five dimensions – plausibility, factuality, typicality, narrative consistency, and perceptual quality. They have also developed their own Perceived Reality Scale for assessing the perceived reality of advertisements on a seven-point Likert scale. In order to capture the multidimensional nature of perceived realism, Cho et al.'s scale was adapted for use in Pilot Study 2. The scale includes five dimensions: *Plausibility* is the key factor in perceived realism and deals with whether the events, settings, and characters presented in media *could* occur in the real-world now or in the future. *Factual realism* is defined as the extent to which the story or narrative represents real-life events and people. *Typicality* refers to the extent to which individuals find the story, characters, and settings similar to their own life and daily experiences. *Narrative consistency* deals with individuals' perception of the coherence of the presented story. Finally, *perceptual quality* refers to the extent to which the auditory and visual qualities of the media create a convincing narrative (or 'reality') within the story itself.

4.3 Pilot Study 2 - Manipulating the Perceived Plausibility, Factual Realism, and Typicality of Robots

Following the lack of success in manipulating participants' belief in the fictional status of two different robots, Pilot Study 2 aimed to address the limitations identified in the previous section. Care was taken to select a new video that depicted largely unknown robots and could be edited in such a way as to retain its narrative structure, contain no or little dialogue, and remain somewhat ambiguous in format as to not bias participants. In terms of ensuring that the selected video was relatively ambiguous in terms of the elements most common in fictional and non-fictional visual media, the selection and editing process was guided by Pouliot and Cowen's (2007) list of features (see Table 4.5). The information given to participants was changed in order to make it more convincing and engaging. In order to keep participants naïve to the true aim of the study, participants were told that the video was either going to be shown at a well-known Documentary and Film festival or an Arts festival, both of which were real events taking place in Sheffield. Additional emphasis was put on the fictionality of the videos and robots by presenting participants with tailored information sheets referring to either the Sheffield Doc/Fest or the Festival of Arts and Humanities. Cho et al.'s adapted Perceived Reality Scale was used to assess whether the manipulation was successful.

Table 4.5

Conventional Visual and Auditory Features of Narrative Fiction and Documentaries

Visual Features	
Narrative Fiction	**Documentary**
• Close shots • Rapid pace of editing • Frequent traveling or moving camera • High spatial and temporal continuity between shots • Subjective editing (e.g., point-of-view shots, flashbacks, etc)	• Long shots • Slow pace of editing • Seldom traveling camera or immobile camera • Discontinuity or low spatial and temporal continuity between shots • Objective editing
Auditory Features	
Narrative Fiction	**Documentary**
• Studio created sounds • Dramatic music • Sound predominantly synchronous • Predominance of indirect verbal address	• Location sounds • Background noises • Sound frequently non-synchronous; voice-overs • Predominance of direct verbal address

Note. Taken from Pouliot and Cowen (2007).

4.3.1 Method

4.3.1.1 Participants

A total of 65 participants completed Pilot Study 2.[10] Thirty-seven members of staff were recruited through the volunteers list for staff at the University of Sheffield and an additional 28 Psychology undergraduates were recruited via the Department of Psychology's Online Research Participation System (ORPS). No other demographic information was collected. Staff members were entered into a draw for a £25 Amazon voucher and undergraduate students received 2 course credits as compensation for their time.

4.3.1.2 Procedure and Materials

Data was collected online via Qualtrics and participants were told that the aim of the study was to investigate people's opinions of fictional or non-fictional media dealing

[10] A priori power analysis was not conducted for this study due to an oversight by the author.

with imagining how technology could shape the future of human society. Participants were randomly allocated to one of two conditions and received slightly different information sheets depending on their assignment. Participants in the ‘non-fictional condition’ (N = 30) received information implying that the video (and robots) that they would see were part of a documentary to be shown at the next Sheffield Doc/Fest. Participants in the ‘fictional condition’ (N = 35) received information implying that the video was part of a short fictional film for the Festival of Arts and Humanities. Prior to watching the video, participants were provided with a short description of the respective festival and the video to further reinforce the status of the robots (i.e., as fictional or non-fictional, see Table 4.6).

Video. Participants watched a fragment of the Official Video for the Robot-Era Project (2014), which was created as part of an EU funded project focusing on developing and implementing robotic technology with the aim to improve the quality of life and care of elderly people. While the video is publicly available on YouTube, it has comparatively low viewership and likes, indicating that it was unlikely to be familiar to participants. The video was edited in such a way as to contain approximately seven minutes of robots performing daily tasks (e.g., grocery shopping) for elderly individuals. Human-robot interaction was also depicted. There was little to no dialogue; however, there was background music which was muted in order to avoid any unintended effects.

Table 4.6

Key Information Given to Participants in the Fictional, Non-Fictional, and Control Conditions

Condition	Information
Fictional	**About the Festival of Arts and Humanities** The Festival of Arts and Humanities is an annual celebration of people, cultures, and art. Hosted by the Faculty of Arts and Humanities at the University of Sheffield, the festival brings a mix of concerts, talks, workshops, and screenings to University staff and students as well as to the general public. **About the video** The video that you are about to watch is from a short fiction film created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the next festival as part of a series of talks, shows, and screenings dealing with imagining how technology could shape the future of human society.
Non-fictional	**About the Sheffield Doc/Fest** Sheffield Doc/Fest is an annual documentary festival. Hosted by documentary makers and academics from the University of Sheffield, the festival brings film screenings, sessions, a marketplace, and interactive and virtual reality exhibitions to University staff and students as well as to the general public. **About the video** The video that you are about to watch is from a documentary created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the next festival as part of a series of documentaries dealing with the way that technology will shape human society in the future.

Perceived realism of the video. Participants were then asked to complete a modified version of Cho et al.'s (2014) Perceived Reality Scale (see Table 4.7) while keeping in mind the nature and intended use of the video. Participants were asked to rate the extent to which they agreed with 19 statements on a seven-point scale (1 – *strongly disagree* to 7 – *strongly agree*). The order in which the statements appeared was randomised. All of the subscales of the Perceived Reality Scale had a moderate to high reliability (i.e., internal consistency) as indicated by Cronbach's α values in Table 4.7. *Plausibility*, *Typicality*, *Factuality*, *Perceptual Quality*, and *Narrative Consistency* ratings were obtained by calculating the mean of all items in each subscale.

Additional questions and feedback from participants. Since it was necessary for participants to remain blind to the experimental manipulation, it was important that they had not watched the video prior to the study and were not aware of the Robot-Era Project. Therefore, participants were asked the following questions: "Have you seen some or all of the video you just watched prior to today?" and "Did you know anything about the video or the topic of the video prior to today?". Participants were also asked what they thought the study was about to ensure they were naïve to the experimental manipulation. Based on their responses, it was judged that all participants remained naïve to the true purpose of the study. Therefore, no participants were excluded from the analysis. Participants were then debriefed and given the chance to share any additional comments they had in regards to the study (see Appendix D for debrief procedure). The implications of participants' feedback for the study's methodology were considered and discussed in the next sections.

Table 4.7

Subscales, Items, and Cronbach's α Values for the Perceived Reality Scale

Subscale	Item	Cronbach's α
Plausibility	Real people would not do the things shown in the video. *	.86
	The video showed something that could possibly happen in real life.	
	The events in the video portrayed possible real-life situations.	
	The story in the video could actually happen in real life.	
	Never in real life would what was shown in the video happen. *	
Typicality	Not many people are likely to experience the events portrayed in the video. *	.80
	What happened to the people in the video is what happens to people in real life.	
	The video portrayed events that happen to a lot of people.	
Factuality	The video was based on facts.	.85
	The video showed something that had really happened.	
	What was shown in the video had actually happened.	
Perceptual Quality	I felt that the overall production elements of the video were realistic.	.80
	The visual elements of the video were realistic.	
	The scenes in the video were realistic.	
Narrative Consistency	The video had a coherent narrative.	.75
	The story portrayed in the video was consistent.	
	Parts of the video contradicted each other. *	
	The story portrayed in the video made sense.	
	The events in the video had a logical flow.	

* Denotes reverse-phrased items.

4.3.2 Results

Treatment of the data and evaluation of assumptions for all statistical analyses can be found in Appendix K. Four independent *t*-tests were conducted to determine if there was significant difference in participants' ratings of plausibility, typicality, factuality, and narrative consistency between the fictional and non-fictional conditions. In addition, a Mann-Whitney U tests was conducted to determine whether there was a significant difference in participants' ratings of perceptual quality between the fictional and non-fictional conditions. A Bonferroni correction was applied to account for the multiple comparisons and an adjusted critical p value of .01 was applied.

Plausibility. The extent to which participants found the story depicted in the video to be plausible, as measured by self-reported ratings, did not differ significantly between the fictional (M = 5.08, SD = 1.17) and non-fictional (M = 4.99, SD = 1.10) conditions, t(63) = .33, p = .743, d = 0.08.

Typicality. The extent to which participants found the events depicted in the video to be typical of real life did not differ significantly between the fictional (M = 3.03, SD = 1.27) and non-fictional (M = 3.19, SD = 1.66) conditions, t(53.90) = - .43, p = .668, d = 0.11.

Factuality. The extent to which participants found the story depicted in the video to be based on facts did not differ significantly between the fictional (M = 3.39, SD = 1.37) and non-fictional (M = 3.36, SD = 1.32) conditions, t(63) = .10, p = .917, d = 0.03.

Perceptual Quality. The distribution of ratings of perceptual quality for the fictional and non-fictional conditions were found to be similar in shape as assessed by visual inspection of a population pyramid. As such, it was appropriate to compare the median scores between the two groups for both subscales. The degree to which participants found the visual depiction of the video to be true to real life did not differ significantly between the fictional (Mdn = 5) and non-fictional (Mdn = 5.33) groups, U = 573.50, z = 0.64, p = .520, r = .08.

Narrative Consistency. The extent to which participants found the story depicted in the video to be coherent and without contradictions did not differ significantly between the fictional (M = 5.44, SD = 0.60) and non-fictional (M = 5.64, SD = 0.80) conditions, t(63) = -1.11, p = .270, d = 0.27.

4.3.3 Discussion

Despite the changes to the procedure and measures, Pilot Study 2 did not find any significant differences between the fictional and non-fictional conditions on any of the five dimensions of perceived realism. As such, the methodology used in Pilot Study 2 was insufficient for manipulating participants' beliefs with respect to the status (fictional or non-fictional) of the video that they were asked to watch and therefore also failed to manipulate their beliefs with respect to the robots depicted in the video. However, this could once again be due to a number of limitations with the methodology and materials used in this pilot study. Not least, that the expected effect is likely to be quite small and as such would require data from a much larger sample than the one recruited for Pilot Study 2 to detect it. For example, in order to detect a small difference (d = .30) between the two conditions with a power of .80 at the .05 significance level using an independent t-test, N = 139 participants per condition would be required.

4.3.3.1 Limitations of the instructions

Successfully manipulating whether participants believed that the video and robots were fictional or non-fictional depended on participants reading and thinking about the information that was displayed prior to the video. While trying to make the manipulation subtler and thus less likely to be detected (and potentially resisted) by participants, it is possible that it was too implicit and thus had no effect. However, it is difficult to say whether the information provided was not convincing enough or if participants did not pay enough attention to it. To investigate this further, questions assessing the extent to which participants paid attention to the information they were given are necessary. Additionally, there was no assessment of whether participants paid attention to the video which could have also affected the success of the manipulation. The relatively low average ratings of typicality and plausibility suggest that participants did not find the video typical of real life. Although this was expected, it also begs the question of whether 'telling' participants that a particular robot is 'real' is the best way to manipulate fictionality. It is possible that any video depicting robots performing daily tasks will be so far removed from most people's real life experiences, that convincing them this is something that is happening in the present time, might be difficult.

4.3.3.2 **Limitations of the measures**

Perceived realism is generally defined as an individuals' subjective evaluation of the extent to which they find a narrative (e.g., a film, novel, advertisement, etc.) reflective of the real world (Cho et al., 2014). However, this should not be confused with what is commonly referred to as *factual* realism – which is whether something has truly occurred in the real world (Green, 2014). The distinction between perceived and factual realism can be problematic as a piece of fiction can be understood to be such yet still be perceived as realistic. Since the aim of Pilot Study 2 was to convince participants that the robots depicted in the video are either fictional or real creations, rather than measure how realistic participants find them to be, in retrospect, the Perceived Reality Scale (Cho et al., 2014) may not have been the ideal measure. Indeed, only one subscale – factuality – measures whether participants believe that the video is based on factual information. However, regardless of the condition to which participants were assigned, the average scores for both groups were around 3 (on a scale from 1 to 7), indicating that participants were generally not convinced that the video depicted events that had taken place in real life. This suggests that the manipulation was insufficient to convince participants that the video and robots were non-fictional.

Additionally, and perhaps most importantly, the Perceived Reality Scale is designed to measure the perceived realism of the video and its narrative rather than the perceived realism of the robots depicted in the video. Whilst some questions refer to the plausibility of the depicted events in general, none directly measure participants' beliefs about the robots. As such, it is impossible to say for sure whether manipulating the fictionality of the robots was successful or not, and to what extent the perceived realism of the robots affected participants' overall evaluation of the video's realism.

Considering the issues outlined above, it was not possible to conclude whether the method used to manipulate the perceived fictionality of the robots was unsuccessful and it was decided that a final pilot study addressing some the limitations was necessary to assess the viability of the methodology and materials for the main study.

4.4 Pilot Study 3 - Successfully Manipulating the Perceived Factual Realism of Robots

As it was possible to edit the video used in Pilot Study 2, it was used once again. However, more text was added to the condition-specific information in order to strengthen the fictional or non-fictional status of the video and robots. There were also some concerns that the two events described in Pilot Study 2, one a documentary festival and the other an art festival, were too well known and might have prompted some participants to look for more information regarding the videos. In order to prevent any unintended effects of familiarity with the events, generic substitutes (i.e., a robotics conference and a sci-fi convention) were used instead. There was also some concern that participants in the first and second pilot studies did not pay attention to the key information and video. In order to address this, attention check measures were included in Pilot Study 3 in order to assess whether participants read the key information and watched the video. Finally, as established in Pilot Study 1, perceived realism has been defined as a multidimensional construct that cannot be captured by a single item. As such, Cho et al.'s measure plausibility, factuality, typicality, narrative consistency, and perceptual quality was used once again in Pilot Study 3. However, the scale was further adapted to not only measure the perceived realism of the video but also that of the robots in order to address some of the methodological issues of Pilot Study 2.

4.4.1 Method

4.4.1.1 Participants

A priori power analysis indicated that a sample of approximately 350 (N = 175 per condition) would be needed to detect a small effect size (f^2 = .05) with a power of .80 at the .05 significance level (see Section 4.3.3). A sample of 284 participants completed the study. Approximately two-thirds of the participants (N = 196) were recruited via email through the staff and student volunteer lists at the University of Sheffield. As a compensation for their time, participants were entered into a draw for a £25 Amazon voucher. The rest of the participants (N = 88) were first year psychology undergraduates who were recruited via the Department of Psychology's Online Research Participation System (ORPS). Students received two course credits as compensation for their time.

Since concealing the experimental manipulation (i.e., fictional vs. non-fictional robots) was key for the aims of the study, it was important that participants had not watched the video prior to the study or had considerable knowledge of the Era-Project or assistive robotics. To determine whether this was the case, participants who indicated that they had seen the video or part of the video prior to the study (N = 19) were selected and their responses to the open-ended question “If yes, briefly describe how and when you have seen some or all of the video.” were explored. Approximately half of those participants seemed to misunderstand the question (e.g., “I have seen all of the video through this online survey”) or their response was not clear enough to warrant their exclusion (e.g., “In a sci-fi movie”). Those who clearly stated that they had seen the video, either in a previous pilot or otherwise, were excluded from further analyses. In order to identify participants who might have had knowledge of assistive robotics, the responses of participants who indicated that they possessed such knowledge (N = 41) to the open-ended question “If yes, please briefly describe what you knew about the video or topic of the video.” were evaluated. While the vast majority of participants provided answers which were either too vague or did not indicate any specialist knowledge of the video topic or robotics (e.g., how robotics is advancing and how it could help future life), some participants indicated professional or academic knowledge of assistive robotics and were therefore excluded from further analyses. Participants who provided the same response (e.g., 7 = *strongly agree*) to all items in one or both of the Perceived Reality Scales were also excluded. To check whether participants had paid attention to the information that they were provided regarding the fictionality of the video and robots, their responses to the three attention check questions were explored. All participants who failed to answer two or more questions correctly or made no attempt to answer those questions were excluded. Table 4.8 provides a summary of participants excluded from the analyses, with reasons. After exclusions, the data of 222 participants was entered for further analysis.

Table 4.8

Number of Participants Excluded and the Reason for their Exclusion

Reason for exclusion	*N* of excluded participants (%)
Participant had viewed some or all of the video prior to the study	10 (3.52)
Participant had indicated considerable knowledge on the topic of the video or robotics	7 (2.46)
Participant had given the same response to all items of one or both self-report scales	3 (1.06)
Participant had failed to answer two or more attention check questions correctly	42 (14.79)
Participant found to be a multivariate outlier	2 (0.70)
Total excluded	64 (22.54)

4.4.1.2 Procedure and Materials

Data was collected online via Qualtrics. Participants were randomly allocated to one of two conditions (fictional or non-fictional) and received slightly different information sheets depending on their assignment. Approximately half of the sample (*N* = 148) received information implying that the video (and robots) they saw was a part of a documentary to be shown a future UK Robotics conference. The other half (*N* = 136) received information implying that the video was part of a short fictional film to be shown at a UK Sci-fi Festival. Participants were told that the aim of the study was to investigate people’s opinions of media dealing with imagining how technology could shape the future of human society. Prior to watching the video, participants were provided with a short description of the event, video, and the robots, to further reinforce the fictional/non-fictional status of the material (see Table 4.9).

Video. Participants were asked to watch a fragment of the official video for the Robot-Era Project (2014), which was created as part of an EU funded project focusing on developing and implementing robotic technology with the aim of improving the quality of life and care of elderly people. The video was approximately seven minutes long and showed robots performing daily tasks (e.g., grocery shopping) and interacting with elderly individuals. There was little to no dialogue and the video was muted in order to avoid any potential effects of background music and sound.

Perceived realism of the video. Participants were then asked to complete a modified version of Cho et al.'s Perceived Reality Scale (see Section 4.3.1.2 for full details). All of the subscales of the Perceived Reality Scale had a moderate to high reliability (i.e., internal consistency) as indicated by Cronbach's α values in Table 4.10.

Table 4.9

Information Given to Participants in the Fictional and Non-Fictional Conditions

Condition	Information
Fictional	**About the UK Sci-fi Festival** The UK Sci-fi Festival is a celebration of science fiction that this year will be hosted by the University of Sheffield. The festival will bring a mix of talks, exhibitions, and screenings to University staff and students as well as to the general public. **About the video** The video that you are about to watch is from a short fiction film created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the festival as part of a series of talks, shows, and screenings dealing with how technology could shape the future of human society. **About the technology** The technology depicted in the video that you are about to see is fictional – that is, it is a story about the way that robots *could* assist us in the future. It does not reflect current reality and the technology that is available to us now, but rather, it is the Robot-Era team’s vision of what robots *could* do in the future. In this video, the Robot-Era team have used a mixture of clever cinematography and remote-controlled props to give life to their imaginary robots.
Non-fictional	**About the UK Robotics Conference** UK Robotics is a conference that this year will be hosted by Sheffield Robotics. The conference will include a showcase of the latest research in robotics across the UK as well as a robot exhibition open to the general public. **About the video** The video that you are about to watch is from a documentary created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the conference as part of a public exhibition dealing with the way that technology shapes human society. **About the technology** The technology depicted in the video that you are about to see is real – that is, it shows robots that have been developed to assist people. These robots are a good representation of existing technology that is available to us now. In this video, the Robot-Era team have filmed some of their robots in action to test and showcase their abilities.

Table 4.10

Subscales Used to Measure the Perceived Realism of the Robot-Era Project Video and Their Reliability as Indicated by Cronbach's α

Subscale	Cronbach's α
Plausibility	.78
Typicality	.76
Factuality	.88
Perceptual Quality	.90
Narrative Consistency	.86*

* Final Cronbach's α after the deletion of an item.

Perceived realism of the robots. The Perceived Reality Scale used to assess the video (Cho et al., 2014) was designed to measure the perceived realism of a video and its narrative, rather than the perceived realism of the actors (in this case, robots and elderly individuals) depicted in the video. Whilst some questions refer to the plausibility of the depicted events in general, none directly measure participants' beliefs about the robots. This was identified as a limitation of Pilot Study 2. As such, in Pilot Study 3, participants were also asked to rate the extent to which they agree with 11 additional statements, relating to the perceived realism of the robots depicted in the video, on a seven-point scale (1 – *strongly disagree* to 7 – *strongly agree*). The additional statements (see Table 4.11) were based on four of Cho et al.'s five subscales (namely, Plausibility, Typicality, Factuality, and Perceptual Quality). The subscale measuring narrative consistency was not included as this subscale only makes sense in relation to a narrative of a story (in this case, the video) rather than specific elements of the narrative. The Factuality and Perceptual Quality subscales both had a relatively high reliability (i.e., internal consistency) as indicated by Cronbach's α values in Table 4.11. Plausibility, Typicality, Factuality, and Perceptual Quality ratings were obtained by calculating the mean of all items in each subscale.

Attention check measures. Participants were asked to answer a few questions about what they have read and seen during the study in order to assess whether they have paid attention to the key information and the video. They were asked the following open-ended questions: "What is the 'UK Sci-fi Festival' (or 'UK Robotics Conference')?", “What was the purpose or topic of the video?”, “How many different robots made an appearance in the video?”, and “What were the robots in the video doing?”. It should be noted that the question regarding the number of robots in the video was disregarded as it was not clear whether participants should indicate the number of individual robots that made an appearance or the number of different models or robots. They were then asked to state whether they have watched the video before and if they knew anything about the topic of the video. Finally, participants were debriefed and asked to leave any comments or suggestions that they might have had as a result of completing the study (see Appendix D for debrief procedure).

Table 4.11

Subscales Used to Measure the Perceived Realism of the Robot-Era Project Robots and Their Reliability as Indicated by Cronbach's α

Subscale	Item	Cronbach's α
Plausibility	Real robots would not do the things shown in the video. [a]	.64*
	The video showed how robots function in real life. [b]	
	The robots shown in the video will never be used in real life. [a]	
Typicality	Not many people are likely to encounter the robots depicted in the video. [a b]	.76*
	The robots depicted in the video are like the robots that people encounter in real life.	
	The video portrayed robots that people often use in real life.	
Factuality	The robots’ abilities were based on the abilities of currently available technology.	.80
	The video showed what robots can do now.	
	The robots shown in the video exist in real life.	
Perceptual Quality	I felt that the robots in the video looked real.	.85
	The robots in the video were realistic.	

[a] Denotes reverse-phrased items.

[b] Denotes deleted items.

* Final Cronbach's α after the deletion of an item.

4.4.2 Results

4.4.2.1 Treatment of the data

To check for univariate outliers, participants' responses for each item (for each of the two conditions) were transformed into *z*-scores and 74 (1.11%) potential outliers, indicated by $2.58 < z < -2.58$ values, were found. Only six (0.09%) of the potential outliers were found to be considerably lower or higher (i.e., a difference equal to or greater than 3) than the participant's mean score for the subscale to which those items belonged. These responses were therefore removed from the data set. To check for multivariate outliers, Mahalanobis distance was calculated for each participant. Two (1.11%) participants were found to be multivariate outliers ($p < .001$), indicated by a critical value > 27.88 for DV = 9 and were excluded from further analysis (see Table 4.8). The final number of participants entered for analysis was 220. An evaluation of the assumptions for a one-way MANOVA can be found in Appendix L.

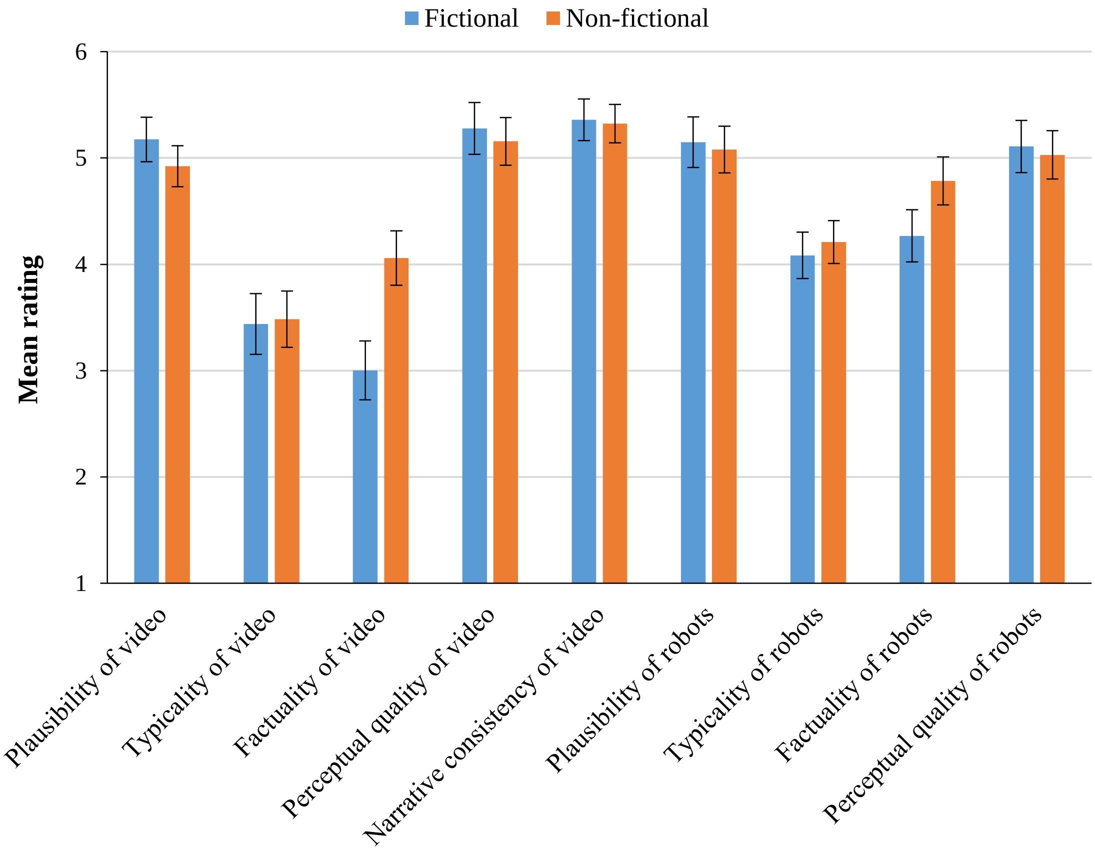
4.4.2.2 Effects of the experimental manipulation on participants' ratings of the video and robot

A one-way MANOVA was conducted to determine whether presenting participants with information that implied that the video and robots that they viewed were fictional or non-fictional, had any effect on how realistic they rated the video and robots to be. A significant effect of the experimental manipulation on participants' ratings of the video and robots was found using Pillai's trace, $V = 0.18$, $F(9, 210) = 5.09$, $p < .001$, $\eta^2_p = .18$, 90% CI [0.08, 0.22].

Separate univariate ANOVAs on the dependant variables revealed a significant effect of the experimental condition (fictional vs. non-fictional) on the Factuality ratings for both the video and the robots respectively, $F(1, 218) = 30.53$, $p < .001$, $\eta^2_p = .12$, 90% CI [0.06, 0.19] and $F(1, 218) = 9.40$, $p = .002$, $\eta^2_p = .04$, 90% CI [0.01, 0.09]. On average, participants who read the information implying that the video was non-fictional ($M = 4.06$, $SD = 0.12$), agreed that the video was based on factual information more than the participants who read the information implying that the video was fictional ($M = 3.00$, $SD = 0.15$; see Figure 4.1). Similarly, participants who read the information implying that the robots in the video were non-fictional ($M = 4.79$, $SD = 0.10$), agreed that the robots' abilities were based on factual information more than the participants who read the information implying that the robots in the video were fictional (*M* = 4.27, *SD* = 0.14; see Figure 4.1).

Figure 4.1

Mean Ratings for Each Subscale of the Adapted Perceived Reality Scales for Each Condition



Note. Error Bars Represent 95% CI.

4.4.2.3 Additional comments by participants

At the end of the study, participants were asked to leave any comments or suggestions that they might have had as a result of completing the study. Thirty-three of the 284 participants left a comment. These comments were read through and assigned one or more labels describing the content or theme of the comment. Below is a breakdown of the major themes that comments fell under.

Sound of the video. Eight participants left comments regarding the lack of sound (audio) of the video. Some participants specified that the video would be better with sound and that sound would have contributed to their understanding of the video narrative. Other participants thought that the lack of sound was a technical problem despite the fact that they were told that the sound of the video was intentionally muted. This is perhaps an indication of participants not paying attention to the information they were provided with during the study.

Quality of the video. Six participants commented on the poor quality of the video. Although the original video is of good quality, some participants clearly experienced issues such as blurring or a delayed playback. This variability in the quality of the video is most likely due to the speed of participants' internet connection or their viewing device's resolution (although, this is less likely if participants completed the study on a PC or laptop).

Measures. Six participants left comments pertaining to the measures (or more specifically the items of the two scales). Some participants stated that they found the questions repetitive while others thought that some questions were confusing and difficult to answer (in particular, questions relating to how 'real' participants found the robots to be).

Experimental manipulation. Three participants commented on the experimental manipulation and gave suggestion for improving the design of the study which mostly involved making changes to the descriptive text implying that the video and robots were either fictional or non-fictional.

General comment about robots/robotics/video. Ten participants left comments relating to their opinion of the robots shown in the video, robotics in general, or the video itself (e.g., "The video was genuinely interesting…"). These comments, while interesting, did not provide additional insight on whether the experimental manipulation was successful.

Other. Five participants made comments which could not be categorised as any of the above mentioned themes. These comments varied but were ultimately not thought to be relevant to the aims or quality of the study.

4.4.3 Discussion

Pilot Study 3 addressed the limitations of previous studies and the results suggest that the experimental manipulation was successful in significantly affecting one, and arguably the most important, aspect of the video and robots' perceived realism, *factuality*.

4.4.3.1 Factuality of the video and robots

According to Cho et al., factual realism (i.e., factuality) is defined as the extent to which the story or narrative represents real-life events and people. Pouliot and Cowen point out that the existing literature suggests that evaluations of factual realism are extremely important when individuals need to distinguishing between a work of fiction (e.g., movie) and non-fictional media (e.g., documentary). As such, we expected participants to rate the video and robots as significantly more 'factual' when the video was described as a part of a documentary in comparison to when described as a work of fiction. The results from the current study support this hypothesis as participants rated both the video and the robots as more factual (i.e., higher mean rating on the factuality subscale) when they were presented with information implying that the video and robots were non-fictional. This finding suggests that presenting participants with different descriptive information prior to watching the video was successful in convincing them that the video and robots were either a work of fiction or non-fiction; or that, at the very least, the video and robots reflected real-life to a different extent.

However, it should be noted that, on average, participants in the non-fictional condition did not rate the video as particularly factual (i.e., based on real-life events and information). We would expect that if participants in the non-fictional condition believed the video to be a documentary (i.e., non-fictional), as described in the information they received, they would have, on average, rated the video as considerably more fictional than they did. One possible explanation is that the items used to measure the construct of factuality as defined by Cho et al. were not valid despite them being based on the original items very closely. However, it is difficult to establish whether this was the case using the data collected in this pilot study.

If we assume that the items we used were indeed measuring factuality, then the relatively low average rating of the video's factual realism could be a product of the scenario depicted in the video. For example, once participant mentioned that despite them being in the non-fictional condition, the video still appeared 'unreal' to them due to their beliefs about the abilities of current robots. Perhaps the extent to which participants can be convinced that the video is non-fictional is moderated by their prior knowledge of robotics or their assumptions regarding the availability and capabilities of existing robots. While it is impossible to say whether this was the case, the possibility that some participants would be more easily convinced that the video in question is fictional rather than non-fictional should be considered when deciding whether the experimental manipulation was successful or not.

4.4.3.2 Plausibility of the video and robots

Plausibility is a key factor in perceived realism and deals with whether the events, settings, and characters presented in the media could occur in the real-world now or in the future (Cho et al., 2014). Plausibility is not limited by the fictional status of the narrative. For example, a work of fiction portraying plausible events could be rated as equally plausible as a documentary. We expected that the plausibility ratings for both the video and the robots would have been significantly higher in the non-fictional condition than in the fictional condition. This hypothesis was not supported by the findings as there was no significant difference between the conditions. While this is not what we expected, in retrospect these results are not very surprising. Since participants in both conditions viewed the same video and the scenes depicted in the video were unlikely to reflect participants' real-life experiences, it is likely that regardless of the information participants received, they found the likelihood of robots performing daily tasks for the elderly in the future equally plausible. On average, participants in both conditions rated the video and robots as relatively high in plausibility which is good as less plausible material is also likely to be less believable as a piece of non-fictional work.

4.4.3.3 Typicality of the video and robots

Typicality refers to the extent to which individuals find the story, characters, and settings similar to their own life and daily experiences (Cho et al., 2014). Since the video participants viewed depicted robots and their use in elderly care and assistance, something that is currently 'not typical', we expected the video to be rated consistently low on typicality features regardless of whether it was implied that the video was fictional or non-fictional. The results from this study supported this hypothesis as there was no significant difference between the conditions and, on average, participants rated both the video and the robot relatively low on typicality.

4.4.3.4 Narrative consistency of the video

Narrative consistency deals with individuals' perception of the coherence of the presented story (Cho et al., 2014). Similar to plausibility, this construct can be independent of the fictional status of the narrative. Fictional work can be equally or more consistent than a piece of non-fictional narrative depending on the way it has been constructed or presented. Since participants watched the same video, we did not expect to see a difference in the narrative consistency ratings participants in either condition gave the video. This was indeed the case and on average participants rated the video as relatively high in narrative consistency. It was important that narrative consistency was not low as previous studies have indicated that poor narrative quality (i.e., a story that does not make sense) can impact the audience's perceptions and opinions of the medium.

4.4.3.5 Perceptual quality of the video and robots

Perceptual quality refers to the extent to which the auditory and visual qualities of the media create a convincing narrative (or 'reality') within the story itself. This construct can be independent from the fictional status of the media. For example, a movie may not reflect the viewer's everyday reality, but can still create a separate and convening 'fictional reality'. This is somewhat related to the video quality measure utilised in the first pilot study. Since we know that the quality of the media affects elements such as transportation and engagement with the narrative (Green, 2014), it was important that participants did not perceive the quality of the video as low. On average, participants rated both the video and the robots as relatively high in perceptual quality although the lack of sound in the video (see below) could have affected these ratings negatively. We predicted that there would be no significant difference between the perceptual quality ratings participants in the fictional and non-fictional condition gave to the video and robots as all participants saw the same video. This hypothesis was supported by the findings.

4.4.3.6 Sample size, power, and implications

It should be noted that the observed effect size of the global effect of the experimental manipulation on participants' ratings of the video and robot was moderate rather than the expected small effect size. Although the expectations regarding effect size were based on a conservative estimate given the lack of literature on the topic, it may still be worth considering what the implications would be for this and future studies. For example, if the power analysis described in section 4.4.1.1 was conducted with a predicted moderate effect size, the achieved sample would have meant that Pilot Study 3 had over 90% power to detect a moderate to large effect using a MANOVA and less than 80% power to detect a small effect with the same test and obtained sample size. The same would be true for the follow-up ANOVA tests. This would then take Pilot Study 3 from what would be considered an "underpowered" study (i.e., higher risk of biased conclusions) to a potentially "overpowered" one (i.e., wasted resources). Therefore, the expected effect size plays a crucial role when conducting priori power analyses and may need to be adjusted for Study 5. However, just because there was a difference of a moderate effect size between the two conditions it terms of the video's *factuality*, it does not necessarily mean that there will be an observable effect of the same magnitude on people's attitudes should the same method be used. As such, where priori power analysis for Study 5 is concerned, the predicted effect size of the difference between people's attitudes should remain as a conservative estimate (i.e., small effect size). The broader implications of power and effect size are discussed in Chapter 6 (Section 6.5).

4.4.3.7 Limitations

It is unclear how many participants experienced issues with the quality or playback of the video and whether such issues would have affected their responses. Literature suggests that the quality of visual material can have an effect on the extent to which people pay attention to the material and engage with the material content (Dobrian et al., 2011). Dealing with the issues participants reported in regard to the quality of the video within Qualtrics could be somewhat tricky as Qualtrics relies on an internet connection to play videos. One possible solution could be to make sure that the video is played at the lowest possible resolution so that all participants experience the same quality. However, video pixilation could be quite distracting even if participants are pre-warned of the poor quality of the video and this could interfere with participants' ability to pay attention to the video. Another solution could be to retain the current setting of the video (i.e., playback at the highest possible resolution given the internet connection) and ask participants to report on the quality of the video via an open-ended question or multiple Likert scales. The quality of the video can then be analysed and accounted for.

The lack of sound was mentioned by participant as a potential issue in not only this pilot study but also in the previous pilot. This could have resulted in participants not paying attention to the video and as such should be reconsidered as the video is fairly long at 7 minutes. One participant also mentioned that the inclusion of the location of the video (i.e., Italy) made them suspicious of the aims of the study as the video description stated that Robot-Era was a Sheffield based project. As such, it is clear that the video may require further editing for the main study to make sure that any clues as to the fictionality of the video and robots are not included in the video itself.

A single item could be added as a final measure of the extent to which participants believed the video and robots were fictional or non-fictional. Since some participants commented that their ratings on the Perceived Reality Scale did not necessarily reflect their perception of the video's fictionality, an item such as "I believed the video/robots to be fictional/non-fictional", rated on a five-point scale (*Agree – Disagree*) could be used to assess whether this was the case.

4.5 Conclusion

Chapter 4 presented three pilot studies that aimed to inform the materials and design used in Study 5. The pilot studies explored how labelling media as fact or fiction (in this case videos depicting robots) could influence individuals' perception of the fictional status of both the robots and videos. Initially, it was expected that it would be fairly easy to influence individuals' perception given that participants were not, generally speaking, familiar with robotics and as such would not be able to accurately differentiate between fictional robots and non-fictional robots. However, this was not the case. Pilot Study 1 was particularly flawed in a number of ways that meant that the intended manipulation was not successful. The issues that became apparent with hindsight were however informative in their own way. Most notably, Pilot Study 1 resulted in a closer inspection of the concept of *realism* and its multidimensional - rather than dichotomous - nature (i.e., real vs. not real). More specifically, the experimental work following Pilot Study 1 in Chapter 4 and 5 focused on the manipulation and measurement of *perceived realism* rather than just the unidimensional *factual realism*. Labelling something as fiction or non-fiction is unlikely to be effective in changing participants' perception regardless of their expertise on a given topic as research shows that people process realism in a more thorough and multifaceted way than apparent at first (Mendelson & Papacharissi, 2007; Busselle & Bilandzic, 2008). Pilot Study 1 also prompted a more thorough process of selecting the video material for which fictionality was manipulated as prior research on media realism and anecdotal evidence from participants' informal feedback showed that media can contain different elements of realness regardless of its original purpose or label.

For example, a fictional film can contain elements of realism that could affect individuals' scrutiny of the content and its subsequent impact on their affective and cognitive responses (Green et al., 2006; Appel & Maleckar, 2012). Alternatively, advertisements tend to present a warped version of factual reality that may contain elements of unrealness or exaggerate facts for the purposes of marketing. Although this line of inquiry was not pursued further, it does raise some questions about how robots are presented in non-fictional mass media and whether there is transparency regarding the abilities and usefulness of technology. Additionally, it begs the question of how robotic devices are perceived in fiction where the setting and robots are realistic verses when the setting and robots are far from grounded in reality (e.g., humanoid robots in the far future).

As discussed above, the results from this pilot study indicate that presenting participants with different information regarding the robots and video did have some impact on how realistic participants perceived the material to be. However, it is difficult to establish whether this difference indicated that participants in the two conditions actually believed the video and robots to be non-fictional or fictional. Additionally, participants' attitudes toward robots were not measured so it is not possible to say whether presenting people with descriptive text implying that the Robot-Era video is fictional or non-fictional would have an effect on said attitudes. As such, it was decided that the main study should be conducted with minor changes in order to establish whether the methodology developed so far would result in a difference of attitudes between the conditions.

Chapter 5: Effect of Fictionality on Explicit and Implicit Attitudes Toward Robots

Chapter Summary

Chapter 5 presents two experimental study and one pilot study that examined the effects of perceived fictionality on participants' attitudes toward robots. The method developed in Chapter 4 was used in Study 5 to manipulate participants' beliefs about the fictional nature of a video and the robots depicted therein. Participants who were led to believe that the video and robots were non-fictional in nature (i.e., represented reality) reported more positive explicit attitudes toward the robots in the video and toward robots in general than did participants who were led to believe that the robots were fictional. This finding suggested that fictional and non-fictional depictions of robots may have a different impact on people's attitudes based solely on participants' perception of their fictionality rather than any other distinguishing features of fictional and real robots such as their appearance. Study 6 took a different approach to examining the impact of fictional vs. non-fictional depictions of robots on people's attitudes by priming participants with images of fictional (e.g., C3P0) and non-fictional (e.g., Asimo) robots. Similar to Study 5, participants who viewed images of the non-fictional robots had more positive implicit (but not explicit) attitudes toward robots than participants who were primed with images of fictional robots. The implications of this finding were unclear but may have been confounded by the use of the Implicit Association Task in the measurement of participants' implicit attitudes.

5.1 Introduction

As discussed in Chapter 1 and Chapter 3, research has shown that fiction can influence people's attitudes and opinions (Green et al., 2004), that is to say what people think and feel about particular topics, objects, or groups. For example, people's acceptance of out-groups, most notably ethnic minorities (Johnson et al., 2013), and even people's expectations of medical procedures such as resuscitation (van den Bulck, 2002) can be based on information obtained from fictional media. Fiction may be especially influential if it depicts groups, activities, or phenomena that are not often encountered in real life but that are repeatedly misrepresented or exaggerated in fictional media. Robots fall into this category as they are generally not common in real life but have been the subject of many works of fiction (Hockstein et al., 2007). While plenty of studies make references to the lack of exposure to robots for the majority of the general population and even mention science-fiction as a major source of often unrepresentative of reality information about robotics and its potential societal impact (e.g., Bumby & Dautenhahn, 1999; Scopelliti et al., 2005; Rosén et al., 2018), only a handful of studies have investigated the influence of fictional representations of robots on people's attitudes (e.g., DiSalvo et al., 2002; Riek et al., 2011; Mubin et al., 2015). This chapter describes two experimental studies that investigate the impact of fictional and non-fictional depictions of robots on participants' attitudes.

Given that most people rarely come into contact with more advanced robotics, it is likely that fictional and media representations of robots shape attitudes, acceptance, and expectations of such technology. This can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology (Kriz et al., 2010). Whilst many studies have looked at people's attitudes toward robots in different contexts (Nomura et al., 2006; Bartneck et al., 2007; Broadbent et al., 2010), it is still unclear what role fictional representations of robots play in the formation and measurement of those attitudes.

As such, Study 5 aimed to investigate whether participants' perception of the fictionality of robots affected their attitudes toward specific robots depicted in a video and robots in general. This question was investigated by manipulating the extent to which participants believe that otherwise identical robots depicted in visual media are fictional (i.e., made up) or non-fictional (i.e., real).

Study 6 aimed to provide insight into the same question as Study 5 using an alternative method by investigating whether covertly priming participants with images of fictional or non-fictional robots in an unrelated task produced any changes in participants' attitudes toward robots in general (see Section 5.4).

5.2 Study 5 - The Effect of Perceived Fictionality on Explicit Attitudes Toward Robots

The primary aim of Study 5 was to investigate whether participants' perception of the fictionality of robots depicted in visual media (i.e., a video) are fictional (i.e., made up) or non-fictional (i.e., real) using the methodology developed in Chapter 4. No specific hypotheses were set but it was expected that there would be a significant difference between the conditions for at least one of the measures of attitudes.

A number of changes and additions were made to this study based on observations made during the pilot studies described in Chapter 4. Most notably, a measure of participants' experience with fictional and non-fictional robots was added as knowledge regarding robotics could influence the extent to which participants can be convinced that the video and robots were fictional or non-fictional. Further additions included measures of explicit and implicit attitudes, video quality, and single-items manipulation-check measures. As for the information participants received, it was not possible to make any major changes that were likely to strengthen the manipulation without running additional pilot studies. Since this was not feasible, the way the experimental manipulation was conducted remained the same as the one used in Pilot Study 3. Some visual changes in terms of the way the information was presented were implemented to ensure participants are more likely to read and pay attention to the information. For example, logos representing each of the two events were added in order to attract participants' attention to the key information.

5.2.1 Method

5.2.1.1 Participants

A priori power analysis indicated that a sample of approximately 300 (N = 150 per condition) would be needed to detect a small effect size (f^2 = .05) with a power of .80 at the .05 significance level. Two hundred and thirty participants completed an online study. However, some participants did not complete the study correctly and were excluded (see Table 5.1), resulting in a sample of 136 participants (N = 92 female, N = 43 male, N = 1 not specified; M_{age} = 34.02, SD_{age} = 13.80). Table 5.2 provides the demographic details of the sample split between the two conditions.

Participants were recruited via email through a list of staff and student volunteers at the University of Sheffield. As a compensation for their time, these participants were entered into a draw for a £25 Amazon voucher. First year psychology undergraduates were also recruited to participate via the Department of Psychology's Online Research Participation System (ORPS). These students received three course credits as compensation for their time. Prolific was also used to recruit participants who were paid £5.10 per hour in order to compensate them for their time. Table 12 details the number of participants recruited via each sample source. As the number of participants recruited from Facebook was less than five, Fisher's Exact Test (*2 x 4*) was conducted to see if there were any significant difference between the expected number of participants and the observed number of participants in each condition (Mehta & Patel, 2011). There was no significant difference between conditions (p = .280) and as such it was unlikely that the varied recruitment strategy contributed to any differences of participants' responses between the two conditions.

Table 5.1

Number of Participants Excluded and the Reason for their Exclusion

Reason for exclusion	*N* of excluded participants (%)
Answered all attention check questions incorrectly.	13 (5.65)
Spent less than 6 minutes and 5 seconds viewing the video.	11 (4.78)
Spent less than 18 seconds on the page containing the condition-dependent information. *	70 (30.43)
Total excluded	94 (40.87)

Note. The order in which participants were excluded is the order in which the reasons for exclusion are listed. Some participants satisfied two or more of the reasons for exclusions.

* This exclusion criterion was set ad-hoc as the priori criteria (exclude participants who are 2 standard deviations away from the average time spent on the page containing the condition-dependent information) was no longer appropriate due to the range of times participants spent on that page (i.e., 1 second to 52 minutes) and the instructions participants were given; "You can leave this [browser] window open if you wish to take a break without terminating your session.". Ad-hoc criteria of "less than 18 seconds" was based on the number of words in the condition-dependent information (average of 180 across the two conditions) and the number of words a good (but not exceptional) reader can read. On average, a good reader is estimated to be able to read up to 600 words per 60 seconds (with over 60% comprehension; estimates can vary). It was assumed that everyone who took part was a 'good reader', although there is no way to confirm whether that was the case or not.

Table 5.2

Participants' Demographic Characteristics

Characteristic	Fictional condition (*N* = 64)	Non-fictional condition (*N* = 72)
Age		
M (SD)	34.14 (14.02)	33.92 (13.67)
Range	18 - 71	18 - 67
Gender Identity		
Female (%)	45 (70.31)	47 (65.28)
Male (%)	18 (28.13)	25 (34.72)
Not specified (%)	1 (1.56)	-
Attained Education (*N*)		
A-levels and below, or equivalent	29	33
Higher National Certificate, Foundation Degree, or equivalent	5	5
Bachelor's degree or equivalent	19	21
Master's degree, Doctoral degree, or equivalent	11	13
Recruitment source		
Prolific	51	48
Volunteer mailing lists	6	9
Facebook	0	2
ORPS	7	13

5.2.1.2 Materials and Procedure

Participants were randomly assigned to one of two conditions automatically using Qualtrics' randomisation function and received slightly different information about the study and the video that they were asked to watch. Approximately one-half of the sample (N = 72) received information implying that the video (and robots) they saw was a part of a documentary to be shown a future UK Robotics conference. The other half of the sample (N = 65) received information implying that the video was part of a short fictional film to be shown at a UK Sci-fi Festival. The true purpose of the study was concealed from participants and they were told that the aim of the study was to investigate people's opinions of media dealing with imagining how technology could shape the future of human society. Prior to watching the video, participants were provided with a short description of the event, video, and the robots, to further reinforce the fictional/non-fictional status of the material. The information that they received was dependant on the condition to which they were assigned (see Table 5.3).

Table 5.3

Key Information Given to Participants in the Fictional and Non-Fictional Conditions

Condition	Information
Fiction	**About the UK Sci-fi Festival** The UK Sci-fi Festival is a celebration of science fiction that this year will be hosted by the University of Sheffield. The festival will bring a mix of talks, exhibitions, and screenings to University staff and students as well as to the general public. **About the video** The video that you are about to watch is from a short fiction film created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the festival as part of a series of talks, shows, and screenings dealing with how technology could shape the future of human society. **About the technology** The technology depicted in the video that you are about to see is fictional – that is, it is a story about the way that robots could assist us in the future. It does not reflect current reality and the technology that is available to us now, but rather, it is the Robot-Era team's vision of what robots could do in the future. In this video, the Robot-Era team have used a mixture of clever cinematography and remote-controlled props to give life to their imaginary robots.
Non-fictional	**About the UK Robotics Conference** UK Robotics is a conference that this year will be hosted by Sheffield Robotics. The conference will include a showcase of the latest research in robotics across the UK as well as a robot exhibition open to the general public. **About the video** The video that you are about to watch is from a documentary created by students at the University of Sheffield working on a project called Robot-Era. The full-length film will be shown at the conference as part of a public exhibition dealing with the way that technology shapes human society. **About the technology** The technology depicted in the video that you are about to see is real – that is, it shows robots that have been developed to assist people. These robots are a good representation of existing technology that is available to us now. In this video, the Robot-Era team have filmed some of their robots in action to test and showcase their abilities.

Participants were also shown two logos with the name of the event above the key information in an attempt to further reinforce the legitimacy of the information (see Figure 5.1). The amount of time that participants spent on the page containing the manipulation-relevant information was covertly recorded and used to identify participants who were unlikely to have engaged with the material (see Table 5.1).

Figure 5.1

Graphics Depicting the Logo for Each Event

(a) (b)

Note. (a) Logo used in the fictional condition; (b) Logo used in the non-fictional condition. Created using royalty free graphics form fotor® (fotor.com).

Video. Following the condition-dependant information, participants were asked to watch a fragment (Robot-Era Project, 2017) of the official video for the Robot-Era Project (2014), which was created as part of an EU funded project focusing on developing and implementing robotic technology with the aim of improving the quality of life and care of elderly people. Three robots from the project appeared in the video: DORO, ORO and CORO (see Appendix M for images of the robots). Participants were instructed to pay attention to the video from start to finish as it was not possible to replay or pause it. The video was 6 minutes and 5 seconds long and showed robots performing daily tasks (e.g., grocery shopping) and interacting with elderly individuals. There was little to no dialogue in the video and the sound was enabled. An alternative link to the video was provided in case participants had trouble with playing the video within Qualtrics. The time that participants spent on the page was covertly recorded and used to identify participants who did not watch the video (either within Qualtrics or via the separate link). Participants who spent less than 6 minutes and 5 seconds on the page without indicating that they watched the video via the alternative link were excluded from further analysis (N = 11, 6.83%).

Attention check for the key information. In order to assess whether participants paid attention to the information presented prior to the video, they were asked three multi-choice questions: (a) "What was the video you just watched a part of?" (e.g., a documentary); (b) "During which upcoming event will the video you just watched be shown in its full length?" (e.g., UK robotics conference); and (c) "What was the technology shown in the video described as?" (e.g., something that is not reflective of reality and the technology that is available to us now). Each question contained four possible answers, two of which were the correct answers for the information given in the fictional and non-fictional condition. Two unrelated but similar potential answers were also included in random order (e.g., "an advertisement" for the first question). The answer to each question was recoded numerically as either correct (1) or incorrect (0) depending on the condition to which each participant was assigned. Participants individual scores were summed to make a single attention score from 0 to 3. Thirteen participants (8.13%) who answered all of the attention check questions incorrectly were excluded from the study.

Anxiety, attitudes toward the use of robots, and perceived enjoyment. Following the attention check, participants were told that the questions they were about to see were related to the robots in the video and were asked to "Imagine that in the future you may need to interact with the robots you saw in the video. Please think about what your feelings, thoughts, and reactions might be if you had to interact with these robots in order to answer the questions.". Photos of the robots were also provided to help participants with this task (see Appendix M).

The questions that participants were given were adapted from Heerink et al.'s (2010) Almere Model Questionnaire which was developed for the measuring of acceptance of socially assistive robots (e.g., robots like the ones shown in the Project-Era video). Unlike the other measure of explicit attitudes in this study, Negative Attitudes toward Robots Scale (NARS), the purpose of this particular set of questions is to assess whether the experimental manipulation had an effect on participants' attitudes toward the robots shown in the video rather than toward robots in general. Only three subscales were used for this study as they were deemed most relevant in regards to the measurement of attitudes toward robots (see Table 5.4). The adapted questionnaire consists of 12 items on a five-point scale (1 - *strongly disagree*, 2 – *disagree*, 3 – *undecided*, 4 – *agree*, 5 – *strongly agree*). The order in which the items were presented was randomised. A mean score, between 1 and 5, for each subscale was calculated by averaging the responses each participant provided for all of the items in that subscale. For the anxiety subscale, a higher score indicates that participants have more anxiety toward robots. A higher mean score for the attitudes toward using the robots indicates more positive attitudes and a higher score for perceived enjoyment also indicates more positives attitudes.

Table 5.4

Items and Cronbach's α Values for Each Subscale of the Almere Model Questionnaire

Subscale	Items	Cronbach's α
ANX: Anxiety (affect)	I would be afraid to make mistakes with the robots. [a]	.79
	I would find the robots intimidating. [a]	
	I would be afraid to break something.	
	I would find the robots scary.	
ATT: Attitude toward using the robots (cognitive)	I think that it’s a good idea to use the robots.	.85
	The robots would make life more interesting.	
	It would be good to make use of the robots. [a]	
PENJ: Perceived Enjoyment (affect)	I would find the robots enjoyable. [a]	.89
	I would find the robots boring. [a]	
	I would find the robots fascinating. [a]	
	I would enjoy the robots talking to me. [a]	
	I would enjoy doing things with the robots. [a]	

Note. Cronbach α was calculated after the exclusions detailed in Table 11 were made.

[a] Denotes items that have been adapted into a ‘would’ statement.

Implicit attitudes toward robots. After completing the above mentioned questionnaire, participants were asked to complete an IAT (Greenwald et al., 1998) – a computer task that has been developed to measure the strength of associations between different pairs of concepts. For the purposes of this study, MacDorman et al.'s (2009) robot-human IAT was adapted following the guidance provided by Greenwald et al. (2003) to assess participants' implicit attitudes toward robots, relative to humans. This particular version of the IAT consisted of seven testing blocks. See Section 3.3.1 for full details about the task.

Explicit general attitudes toward robots. Following the IAT, participants were asked to complete a modified version of the Negative Attitudes toward Robots Scale (NARS) to measure their explicit attitudes toward robots (Nomura et al., 2004). The NARS is comprised of three subscales (Interaction with robots, Social influence of robots, and Emotion in interaction with robots) with a total of 16 items (see Table 5.5). Each item is a statement that can be rated on a five-point scale (1 - *strongly disagree*, 2 – *disagree*, 3 – *undecided*, 4 – *agree*, 5 – *strongly agree*). The order in which the items were presented was randomised. The reliability and validity of the NARS has been supported by multiple studies (Nomura et al., 2004; Nomura et al., 2006).

Participants' mean for each of the three subscales of the NARS was generated by summing up the score for each of the items in the Interaction, Social influence, and Emotion subscales as per the instructions in Nomura et al. (2004). The scores of the Emotion subscale were inverted prior to summation. For the Interaction subscale, the minimum score is 6 and the maximum score is 30; for Social influence the minimum is 4 and the maximum is 20 (after the removal of an item); and for the Emotion subscale, the minimum is 5 and the maximum is 25. Larger values indicate more negative attitudes, while smaller values indicate more positive attitudes.

An open-ended question was presented to participant immediately afterwards to explore whether participants had any particular robots or experiences in mind whilst completing the NARS ("Did you have any particular robot(s) or experience(s) in mind when you were answering the questions on the previous page?"). This question was intended as a way to gain some additional insight into what representations of robots people may have in mind when completing measures of their attitudes toward robots.

Table 5.5

Items and Cronbach's α Values for Each Subscale of the Negative Attitudes Toward Robots Scale (NARS)

Subscale	Items	Cronbach's α
NARS-S1: Interaction with robots	I would feel uneasy if I was given a job where I had to use robots.	.80
	The word “robot” means nothing to me.	
	I would hate the idea that robots or artificial intelligences were making judgements about things.	
	I would feel nervous operating a robot in front of other people.	
	I would feel nervous just standing in front of a robot.	
	I would feel paranoid talking with a robot.	
NARS-S2: Social influence of robots	I would feel uneasy if robots really had emotions.	.76*
	Something bad might happen if robots developed into living beings.	
	I feel that if I depend on robots too much, then something bad might happen.	
	I am concerned that robots would be a bad influence on children.	
	I feel that, in the future, society will be dominated by robots. [a]	
NARS-S3: Emotion in interaction with robots (inverse)	I would feel relaxed talking with robots.	.86
	If robots had emotions, then I would be able to make friends with them.	
	I feel that I could make friends with robots.	
	I would feel comfortable being with robots.	
	I would feel comforted being with robots that have emotions. [b]	

Note. Cronbach α was calculated after the exclusions detailed in Table 1 were made.

[a] Denotes items that were removed in order to improve the overall reliability of the scale.

[b] Denotes items that have been adapted into a ‘would’ statement.

* Cronbach's α after item deletion.

Manipulation check. Two measures were used in order to assess whether the experimental manipulation was successful in convincing participants that the story and robots presented in the video were fictional or non-fictional.

Perceived realism of the robots. Immediately after completing the first set of self-report measure, participants were asked to rate the perceived realism of the robots. The Perceived Reality Scale used to assess the video (Cho et al., 2014) was designed to measure the perceived realism of a video and its narrative, rather than the perceived realism of the actors (in this case, robots and elderly individuals) depicted in the video. Whilst some questions refer to the plausibility of the depicted events in general, none directly measure participants' beliefs about the robots. This was identified as a limitation in Pilot Study 2. As such, participants were asked to rate the extent to which they agreed with nine statements, relating to the perceived realism of the robots depicted in the video, on a five-point scale (1 - *strongly disagree*, 2 – *disagree*, 3 – *undecided*, 4 – *agree*, 5 – *strongly agree*). The items (see Table 5.6) were based on three of Cho et al.'s five subscales (namely, *Plausibility*, *Typicality*, and *Factuality*). The subscales measuring narrative consistency and perceptual quality were not included as the first subscale only makes sense in relation to a narrative of a story (in this case, the video) rather than specific elements of the narrative and the later was replaced with a video quality questionnaire. Plausibility, typicality, and factuality ratings were obtained by calculating the mean of all items in each subscale. Higher ratings indicated that participants found the robots in the video and/or their abilities to be more plausible, more typical of real life, and based on factual information.

Table 5.6

Items and Cronbach's α Values for Each Subscale of the Perceived Realism of Robots Scale

Subscale	Item	Cronbach's α
Plausibility	Real robots would not do the things shown in the video. [b]	.65[*]
	The video showed how robots function in real life. [a]	
	The robots shown in the video will never be used in real life. [b]	
Typicality	Not many people are likely to encounter the robots depicted in the video. [a b]	.71[*]
	The robots depicted in the video are like the robots that people encounter in real life.	
	The video portrayed robots that people often use in real life.	
Factuality	The robots' abilities were based on the abilities of currently available technology.	.87
	The video showed what robots can do now.	
	The robots shown in the video exist in real life.	

Note. Cronbach α was calculated after the exclusions detailed in Table 1 were made.

[a] Denotes items that were removed in order to improve the overall reliability of the scale.

[b] Denotes reverse-phrased items.

[*] Cronbach's α after item deletion.

Belief in the fictionality of the story and robots. In addition to the perceived realism of the robots, participants were also asked to indicate whether they believed that the robots in the video were fictional (i.e., do not exist in real life) or non-fictional (i.e., exist in real life) while watching the video; and whether they believed that the story depicted in the video was fictional (mostly made-up for the purposes of the video) or non-fictional (a relatively accurate representation of real-life events). The response for each of the two questions was recoded as either matching the condition to which participant were assigned or not matching the condition.

Video quality. Poor playback of the video was identified as a potential issue based on participants' comments from all three pilot studies. Since the quality of video playback largely depends on the individual viewing device and internet connection, participants were asked to evaluate four different aspects of the video on a 5-point scale from 1 (Bad) to 5 (Excellent). Participants were asked to rate the quality of the video's initial buffering time (i.e., how long did it take for the video to start playing), re-buffering (i.e., how often did the video stop and start playing again), audio (i.e., how clear, loud, and in-sync with the video was the sound), and visual content (i.e., how pixelated, unclear, and distorted was the video). Additionally, participants were also asked whether the video started automatically and if they experienced any additional issues. An overall video quality score was calculated by taking the average of the ratings for initial buffering time, re-buffering, audio quality, and visual quality.

Experience with fictional and non-fictional robots. Following the assessment of video quality, participants were asked to report how many robot-related experiences they have had during their lifetime or in the past year (see Table 5.7) on a 10-point scale from 0 experiences to 10 or more experiences. This measure was adapted from the Robot-related Experiences Questionnaire (MacDorman et al., 2009). For each type of experience with robots, the number of experiences that the participant reported for each item were summed so that the resulting value was between 0 – 20 for indirect experiences with non-fictional robots and fictional robots, and between 0 – 30 for direct experiences with robots.

Demographic questions. Finally, participants were asked to provide their age, gender identity, and highest attained educational qualification. They were then debriefed and given the chance to share any additional comments they had in regards to the study (see Appendix D for debrief procedure).

Table 5.7

Items from MacDorman, Vasudevan, and Ho's (2009) Robot-Related Experiences Questionnaire

Type of experience	Item
Experiences with fictional robots	Read stories, comics, or other fictional material about robots. [a]
	Watched movies, fictional TV programmes, or other media about robots. [a]
Experiences with non-fictional robots (indirect)	Read news articles, product descriptions, conference papers, journal papers, or other factual material about robots. [a]
	Watched documentaries, factual TV programmes, or other factual media about robots. [a]
Experiences with non-fictional robots (direct)	Built or programmed a robot. [b]
	Had physical contact with a robot (not including appliances such as robotic vacuum cleaners). [b]
	Attended lectures, exhibitions, trade shows, competitions, or other events related to robots. [b]

[a] Number of activities performed in the past 1 year.

[b] Number of activities performed in one's life so far.

5.2.2 Results

The raw data was extracted from Qualtrics and copied into an Excel workbook. All analyses were conducted using SPSS.[11]

5.2.2.1 Treatment of the data

Evaluation of assumptions. A one-way MANOVA was conducted to investigate whether there was a significant difference in people's attitudes toward robots based on whether they received information implying that the video and robots were fictional or non-fictional. The assumptions tested are based on the reccommendation made by Finch (2005) and French et al. (2008).

To check for univariate outliers, participants' responses to each item (for each of the two conditions) for the anxiety, attitudes toward use of robot, perceived enoyment, and the three NARS subscsles were transformed into *z*-scores.[12] Six (0.13%) outliers, indicated by $3 < z < -3$ values, were found. These responses were therefore removed from the data set prior to any further analysis. Following this, composite scores for each dependant measure was calculated as described in the Materials and Procedure section. The *d*-scores from the IAT were also calculated as described in the Methods section.

To check for multivariate outliers, Mahalanobis distance was calculated for each participant as this is one of the recommended and most commonly used ways of detecting outliers, given that the dependant variables are not strongly correlated (as is the case for this study; de Maesschalck et al., 2000; Warren et al., 2011). No participants were identified as multivariate outliers based on a critical Chi-Square value > 24.32 for DV = 7, at $p < .001$. Data from $N = 136$ participants across the two conditions ($N_{fictional} = 64$, $N_{non\text{-}fictional} = 72$) was analysed.

The Shapiro-Wilk test of normality was significant for two of the dependant variables (anxiety and attitude toward the use of the robots), for one of the conditions (see Table 5.8 for *W* and *p* values), indicating a potentially non-normal distribution. Further

[11] Unless otherwise stated, the treatment of the data and its analysis was conducted following a priori procedure and in the order in which the analyses are presented in this section.

[12] Not including the items removed in order to improve the reliability of the subscales. See Materials section.

examination of Z_s and Z_k values showed that only attitude toward the use of robots was substantially negatively skewed for the non-fictional condition (indicated by $2.58 < z < -2.58$ for $N > 50$). Given that this deviation was only slightly above the cut-off point and that analysis of variance tests are generally considered fairly robust to deviations from normality for relatively large samples (Todorov & Filzmoser, 2010), it was decided to proceed with the analysis without any transformation of the data. Visual inspection of Q-Q plots also confirmed that the distribution of scores for both conditions and all variables was approximately normally distributed. The assumption of homogeneity of regression slopes was met for all covariates ($p \geq .651$). Error variances were homogeneous for all but one of the seven variables as assessed by Levene's Test of Homogeneity of Variance ($p > .05$). Attitude toward the use of the robots was found to have unequal error variances across the two conditions, $p = .029$. While this is not ideal, for relatively large samples and approximately equal group sizes, reporting Pillai's trace for one-way MANOVAs has been recommended as a solution (Finch, 2005) and has subsequently been used in this study. In addition, heterogenious variances for only one of the variables has been cited (Beasley & Sheehan, 1994) as arguably less problematic than heterogenious variances for the majority of variables.

Table 5.8

Descriptive Statistics for Each Dependent Variable and for Each Condition

Measure	Condition	*M* [95% CI]	*SD*	*W*	df	*p*	Z_s	Z_k
Anxiety (ANX)	Fictional	2.77 [2.55, 2.99]	0.93	0.96	64	.035*	-0.40	-1.63
	Non-fictional	2.61 [2.40, 2.82]	0.87	0.97	72	.101	0.34	-0.77
Attitude (ATT)	Fictional	3.24 [3.02, 3.46]	1.00	0.97	64	.127	-1.03	-0.85
	Non-fictional	3.79 [3.58, 4.00]	0.81	0.94	72	.002*	-2.71*	2.04
Perceived enjoyment (PENJ)	Fictional	3.35 [3.13, 3.57]	0.97	0.97	64	.096	-0.98	-0.49
	Non-fictional	3.72 [3.52, 3.93]	0.79	0.97	72	.057	-0.71	-1.24
Implicit attitudes (IAT)	Fictional	0.55 [0.47, 0.64]	0.37	0.98	64	.493	-0.84	-0.42
	Non-fictional	0.52 [0.43, 0.60]	0.33	0.98	72	.320	-1.54	0.14
Interaction with robots (NARS-S1)	Fictional	16.81 [15.75, 17.87]	4.51	0.98	64	.523	0.03	-0.74
	Non-fictional	14.51 [13.52, 15.51]	4.07	0.97	72	.129	1.22	-0.45
Social influence of robots (NARS-S2)	Fictional	14.08 [13.25, 14.90]	3.28	0.97	64	.067	-1.96	0.87
	Non-fictional	12.76 [11.99, 13.54]	3.38	0.98	72	.382	-0.33	-0.41
Emotion in interaction with robots (NARS-S3)	Fictional	16.41 [15.34, 17.47]	4.21	0.98	64	.552	-0.37	-1.09
	Non-fictional	14.28 [3.27, 15.28]	4.41	0.98	72	.199	-0.85	-1.04

Note. Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. *W*, df, and *p* detail the results from the Shapiro-Wilk test of normality.

* Indicates substantial departure from the normal distribution.

Examination of scatterplots for both conditions indicated an approximately linear relationship between the dependant variables. There appeared to be no multicollinearity between the dependant variables as indicated by correlations of $r < .90$ (see Table 5.9 and 5.10). The majority of variables appeared to be significantly correlated with each other and it was judged this was sufficient (although in some cases less than the recommended moderate correlation between DVs; French et al., 2008) to run a MANOVA. However, the one exception was the *d*-scores from the IAT which did not correlate with any of the variables. As such, it was decided that an independent *t*-test would be more appropriate to investigate whether there was any significant difference in participants' implicit attitudes between the two conditions.

Table 5.9

Correlations (Pearson's r) Between the Dependant Variables for the Fictional Condition

Variables	ANX	ATT	PENJ	IAT	NARS-S1	NARS-S2
Anxiety (ANX)	-	-	-	-	-	-
Attitude toward the use of robots (ATT)	-.03	-	-	-	-	-
Perceived enjoyment (PENJ)	-.02	.83**	-	-	-	-
Implicit attitudes (IAT)	-.24	.17	.13	-	-	-
Interaction with robots (NARS-S1)	.49**.	-.41**	-.42**	-.12	-	-
Social influence of robots (NARS-S2)	.32**	-.45**	-.46**	-.10	.56**	-
Emotion in interaction with robots (NARS-S3)	.15	-.72**	-.72**	.02	.57**	.54**

**p* ≤ .05; ** *p* ≤ .01.

Table 5.10

Correlations (Pearson's r) Between the Dependant Variables for the Non-Fictional Condition

Variables	ANX	ATT	PENJ	IAT	NARS-S1	NARS-S2
Anxiety (ANX)	-	-	-	-	-	-
Attitude toward the use of robots (ATT)	-.39**	-	-	-	-	-
Perceived enjoyment (PENJ)	-.32**	.78**	-	-	-	-
Implicit attitudes (IAT)	-.01	.06	.09	-	-	-
Interaction with robots (NARS-S1)	.65**	-.48**	-.46**	.05	-	-
Social influence of robots (NARS-S2)	.42**	-.43**	-.52**	-.08	.55**	-
Emotion in interaction with robots (NARS-S3)	.39**	-.59**	-.63**	.08	.62**	.63**

**p* ≤ .05; ** *p* ≤ .01.

5.2.2.2 Effect of condition on the dependant variables

A one-way MANOVA was conducted with condition as the independent variable and anxiety, attitudes toward the use of robots, perceived enjoyment, and all NARS subscales as the dependant variables. *D*-scores from the IAT were analysed separately using an independent samples *t*-test. A Bonferroni correction was applied to account for multiple post-hoc comparisons, resulting in a new critical value of p = .007 (7 comparisons).

A one-way MANOVA was conducted to investigate whether there was a difference in participants' attitudes toward robots between the fictional and non-fictional condition. The assumption of homogeneity of variance-covariance matrices was met as indicated by Box's test of equality of covariance matrices (p = .661). Results indicated that there was a significant difference between participants' attitudes toward robots between the two conditions, Pillai's Trace V = .12, F(6, 129) = 2.79, p = .014, η^2_p = .12, 90% CI [0.01, 0.17]. Separate univariate ANOVAs were used to follow-up the significant MANOVA.

Explicit attitudes toward the robots depicted in the video.[13] There was no significant difference (mean difference = 0.16, BCa 95% CI [-0.14, 0.47]) in the level of anxiety participants reported toward the robots between the fictional (M = 2.77, SD = 0.93) and non-fictional (M = 0.55, SD = 0.37) conditions, F(1, 134) = 1.13, p = .289, η^2_p = 0.01, 90% CI [0.00, 0.05].

There was a significant difference (mean difference = -0.55, BCa 95% CI [-0.86, -0.24]) between the two conditions in the level of postive attitudes that participants reported toward the robots in the video was significant, F(1, 134) = 12.69, p = .001, η^2_p = 0.09, 90% CI [0.03, 0.17]. These findings indicate that participants in the fictional (M = 3.24, SD = 1.00) condition had significantly less positive attitudes toward the robots in the video than participants in the non-fictional (M = 3.79, SD = 0.81) condition.

[13] In cases where multiple comparisons are decided upon priori (such as in this case), a Bonferroni correction may inflate the risk of Type II error. As such, the mean difference between conditions followed by bootstrapped 95% CIs have been reported for each *t*-test in order to aid the interpretation of the results.

There was no significant difference (mean difference = -0.37, BCa 95% CI [-0.67, -0.07]) between the fictional (M = 3.35, SD = 0.97) and non-fictional (M = 3.72, SD = 0.79) conditions in the level of perceived enjoyment that participants reported after applying the corrected critical p value, $F(1, 134) = 6.07$, $p = .015$, $\eta^2_p = 0.04$, 90% CI [0.00, 0.11].

Explicit attitudes toward robots in general. Results showed that the mean difference (2.30, BCa 95% CI [0.84, 3.75]) between the two conditions for participants' negative attitudes toward interaction with robots (NARS-S1) was significant, $F(1, 134) =$ 9.75, $p = .002$, $\eta^2_p = 0.07$, 90% CI [0.02, 0.14]. These findings indicate that participants in the fictional condition (M =16.81, SD = 4.51) had significantly more negative attitudes toward interacting with robots than participants in the non-fictional (M = 14.51, SD = 4.07) condition.

There was no significant difference (mean difference = 1.31, BCa 95% CI [0.18, 3.76]) between participants' negative attitudes toward the social influence of robots (NARS-S2) for the ficitional (M = 14.08, SD = 3.28) and non-fictional (M = 12.76, SD = 3.38) conditions after aplying the corrected critical p value, $F(1, 134) = 5.26$, $p = .023$, $\eta^2_p = 0.04$, 90% CI [0.00, 0.10].

The mean difference (2.13, BCa 95% CI [0.66, 3.60]) between participants' negative attitudes toward emotion when interacting with robots (NARS-S3) was significant, $F(1, 134) = 8.24$, $p = .005$, $\eta^2_p = 0.06$, 90% CI [0.01, 0.13]. These findings indicate that participants in the fictional condition (M = 16.41, SD = 4.21) had significantly more negative attitudes toward emotion when interacting with robots than participants in the non-fictional condition (M = 14.28, SD = 4.41).

Implicit attitudes toward robots. Results from the independent samples t-test indicated that there was no significant difference between participants' implicit attitudes (as measured via the IAT) for the fictional (M = 0.55, SD = 0.37) and non-fictional (M = 0.52, SD = 0.33) conditions, $t(134) = .62$, $p = .536$, $d = .0.09$. In general, participants in both conditions showed a slight preference for humans over robots as indicated by d-scores > 0.

5.2.2.3 Analysis of the qualitative Negative Attitudes toward Robots Scale (NARS) question

An open-ended question was presented to participant immediately afterward completing the NARS to explore whether participants had any particular robots or experiences in mind whilst completing the questionnaire ("Did you have any particular robot(s) or experience(s) in mind when you were answering the questions on the previous page?"). This question was intended as a way to gain some additional insight into what representations of robots people may have in mind when completing measures of their attitudes toward robots. Only 39 (28.68% of the total sample) participants stated that they had something in mind when answering the NARS. Comments from those participants were coded based on the type of robots they mentioned in the comment (i.e., fictional, non-fictional, both fictional and non-fictional robots, or only robots from the video) and the valence of the comment (i.e., positive, negative, or neutral).

Fisher's Exact Test (*2 x 4*) was conducted to see if there were any significant difference between the expected number and the observed number of types of robots mentioned by participants in each condition (see Table 5.15). There was no significant difference between the two conditions, $p = .147$. Overall, participants mentioned fictional and non-fictional robots a similar number of times. Participants in the non-fictional condition mentioned the robots in the video more often than participants in the fictional condition. However, this was likely due to the way the comments were coded. Only comments that mentioned the robots from the video and no other type of robot were coded under the last category. However, once all mentions of the robots in the video were counted, the number of times they were mentioned was approximately the same for the fictional ($n = 7$) and non-fictional ($n = 9$) condition.

Table 5.11

Number of Comments Relating to Four Different Types of Robots in the Fictional and Non-Fictional Condition

Type of robot	Condition		
	Fictional	Non-fictional	*n*
Fictional	8	5	13
Non-fictional	6	4	10
Both fictional and non-fictional	4	2	6
Robots from the video	2	8	10
Total *n*	20	19	39

Fisher's Exact Test (*2 x 3*) was conducted to see if there were any significant difference between the expected number and the observed number of positive, negative, and neutral comments regarding robots made by participants in each condition (see Table 5.16). There was no significant difference between the two conditions, $p = .999$. Overall, participants' comments were predominantly neutral (i.e., did not indicate any positive or negative thoughts/emotions toward robots) and only one participant's comment was coded as positive. However, this may be a result of the way in which the comments were coded. The comments of participants who expressed both positive and negative feelings/thoughts toward robots were coded as neutral which may have contributed to the overall number of neutral comments.

Table 5.12

Number of Positive, Negative, and Neutral Comments Relating to Robots for the Fictional and Non-Fictional Condition

Valence of comment	Condition		
	Fictional	Non-fictional	*n*
Positive	1	0	1
Negative	6	5	11
Neutral	13	14	27
Total *n*	20	19	39

5.2.2.4 Manipulation check

Plausibility, typicality, and factuality ratings of the robots' abilities. Based on results from the third pilot study, we expected to find a significant difference between the conditions for the extent to which participants believed that the robots' abilities were based on fact, rather than fiction. We did not expect to find differences in how typical participants found the robot was across the conditions and it was expected that, in general, participants would find the robots less typical when compared to real life. As for how plausible participants found the robots to be, we expected that there would be no significant difference between the conditions based on the results of the pilot study.

Shapiro-Wilk test of normality was significant for plausibility, typicality and factuality in both conditions (see Table 5.11), indicating a non-normal distribution of the ratings. Z_s and Z_k values indicated that only Factuality ratings for the non-fictional condition were substantially negatively skewed and leptokurtic ($2.58 < z < -2.58$ for $N > 50$). However, it should be noted that almost all variables were moderately skewed and/or kurtotic, in different ways. Given the inconsistency of deviations from the normal it was decided that a non-parametric test would be more appropriate for identifying any differences between the conditions.

Three Mann-Whitney U tests were conducted to investigate whether there was a significant difference in the distribution of participants' factuality, plausibility, and typicality ratings between the two conditions. Results indicated that participants in the non-fictional condition found the robots' abilities to be moderately more based on factual information (mean rank = 87.72) than participants in the fictional condition (mean rank = 46.88) did, $U = 3688$, $z = 6.08$, $p < .001$, $r = .52$.[14] Participants in the non-fictional condition also found the robots' abilities slightly more plausible (mean rank = 75.19) than participants in the fictional condition (mean rank = 60.97) did, $U = 2786$, $z = 2.17$, $p = .300$, $r = .19$. There was also a significant difference in how typical participants found the robots to be; participants in the non-fictional condition (mean rank = 74.90) rated the robots as slightly more typical than participants in the fictional condition (mean rank = 61.30), $U = 2765$, $z = 2.04$, $p = .041$, $r = .17$. However, when a Bonferroni correction was applied to account for the multiple tests (critical $p_{corrected} = 0.05/3 = .017$), typicality and plausibility no longer met the cut off criteria for significance.

---

[14] Effect size was calculated using the following formula $r = \frac{z}{\sqrt{N}}$ where z is the standardised test statistic and N is the sample size (Fritz et al., 2012); r to be interpreted as 0.1 (small effect), 0.5 (moderate effect), and 0.7 (large effect).

Table 5.13

Descriptive Statistics for the Plausibility, Typicality, and Factuality Ratings

Measure	Condition	*M*	*SD*	*W*	df	*p*	Z_s	Z_k
Factuality	Fictional	2.89	0.96	0.96	64	.019*	-0.74	-1.62
	Non-fictional	3.94	0.80	0.92	72	.001*	-3.81*	3.59*
Plausibility	Fictional	3.68	0.85	0.94	64	.003*	-1.90	0.76
	Non-fictional	3.99	0.64	0.91	72	.001*	-0.95	-0.53
Typicality	Fictional	2.45	0.94	0.94	64	.004*	0.65	-1.62
	Non-fictional	2.78	0.87	0.95	72	.010*	1.02	-0.21

Note. Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis.

* Indicates substantial departure from the normal distribution.

Participants' belief in the fictionality of the video's story and the robots. Approximately half or more of the participants believed that video that they watched and the robots depicted within were fictional or non-fictional in accordance with the condition they were assigned to (see Table 5.12). This implies that, in general, participants believed the information that they were given about the video and the robots. However, it should be noted that a substantial number of participants provided a different response to the one matching their condition assignment.[15]

Table 5.14

Number of Participants who Believed the Robots and Story in the Video to be Fictional or Non-Fictional by Condition

Condition	I believed the video to be…		I believed the robots to be…		
	Fictional	Non-fictional	Fictional	Non-fictional	N
Fictional	41 (64.1%)	23 (35.9%)	30 (46.9%)	34 (53.1%)	64
Non-fictional	38 (52.8%)	34 (47.2%)	21 (29.2%)	51 (70.8%)	72

[15] The data analysis was repeated post-hoc using only the data from participants whose response to this question matched their condition assignment. This analysis can be found in Appendix N.

5.2.2.5 Video quality

On average, participants reported that the video had a good quality (M = 4.63, SD = 0.40) with the lowest reported quality being 3.50 out of 5. There was no significant difference of the reported video quality between participants in the fictional condition (M = 4.65, SD = 0.44) and non-fictional condition (M = 4.62, SD = 0.38), $t(134) = 0.39$, p = .700. In regards to video playback, 71.32% participants reported that the video did not start automatically. This is not surprising as some popular internet browsers (e.g., Google Chrome) do not allow for the automatic playback of video and audio files. Overall, reported video quality appeared to be fairly consistent across participants and conditions, and instructions given to participants on how to start the video should account for issues with automatic playback. Reported video quality was, therefore, not considered further as a potential factor that could influence participants' responses.

5.2.2.6 Participants' experiences with fictional and non-fictional robots

On average, participants had 6.99 (SD = 5.96) experiences relating to fictional robots and 6.29 (SD = 5.14) indirect experiences relating to non-fictional robots in the past year; and 2.31 (SD = 4.51) direct experiences with non-fictional robots in their entire life. It should, however, be noted that participants' number of experiences with fictional and non-fictional robots varied substantially, with some participants having no robot-related experiences while others indicated that they have had 10 or more robot related experiences for each item. Upon examining z-scores for skewness and kurtosis, it was found that the responses for direct experience with non-fictional robots for the fictional (z_s = 13.74, z_k = 36.14) and non-fictional (z_s = 11.15, z_k = 19.21) condition were, unsurprisingly, extremely positively skewed and leptokurtic. The remaining types of experiences with robots were also substantially positively skewed ($2.58 < z < -2.58$ for $N > 50$). Given this deviation from the normal distribution, post hoc non-parametric tests were performed instead of the planned parametric equivalents and the median values rather than the means are reported for each condition and experience in Table 5.13. Three Mann-Whitney U tests were conducted to investigate whether there was a difference in the number of experiences participants had with fictional and non-fictional robots. There were no significant differences in the median number of experiences that participants in the fictional condition had compared to participants in the non-fictional condition ($p \geq$ .592).

Table 5.15

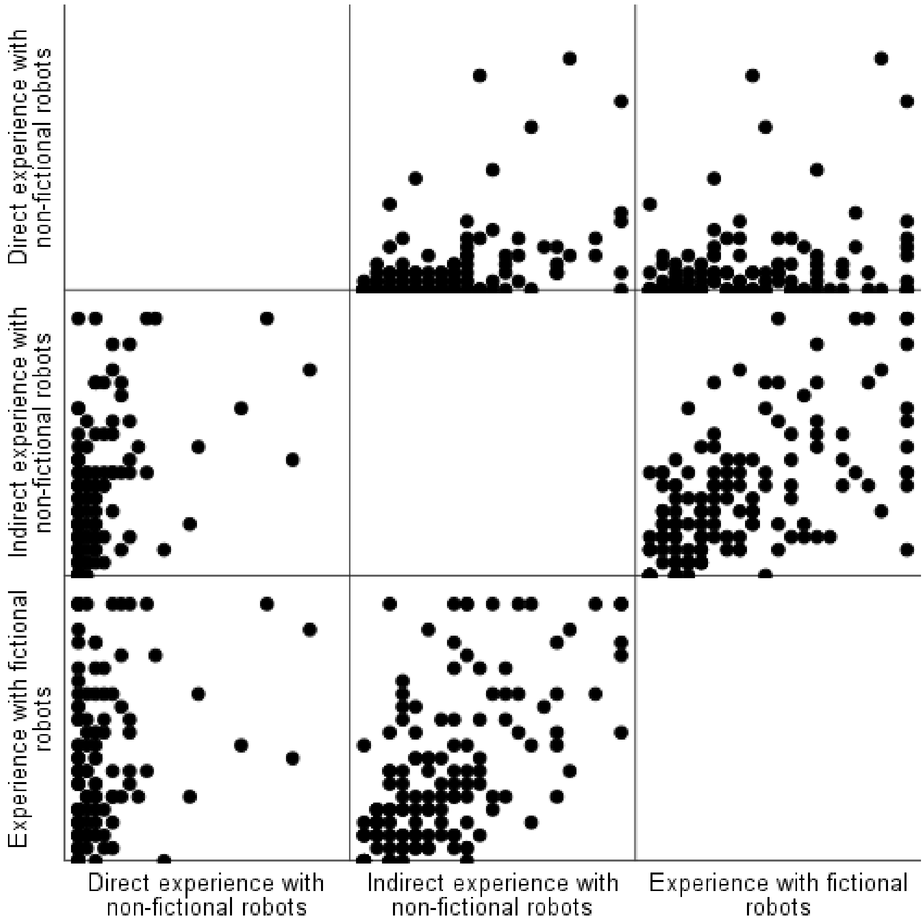
Descriptive Statistics for Experience with Robots by Type of Experience and Condition

Type of experience	Condition	Median	Range
Experience with fictional robots	Fictional	5	0 - 20
	Non-fictional	5.5	0 - 20
Experience with non-fictional robots (indirect)	Fictional	5	0 - 20
	Non-fictional	6	0 - 20
Experience with non-fictional robots (direct)	Fictional	1	0 - 27
	Non-fictional	0.5	0 - 25

Multiple Spearman rho's correlations were planned to be conducted in order to investigate the relationship between the three types of experience with robots. However, upon visual examination of a scatterplot matrix of the three variables (see Figure 5.2), it was found that correlational analysis would not be appropriate for the direct experience with robots due to the high number of 'outliers' (i.e., a few participants who had considerably more experiences than the sample average). Therefore, Spearman rho was carried out only for experience with fictional robots and indirect experience with non-fictional robots. Results showed that there was a strong positive correlation between the two types of experiences (r = .60, p < .001) indicating that, in general, participants who reported a high number of experiences related to fictional robots also reported a high number of indirect experiences with non-fictional robots.

Figure 5.2

Scatterplot Matrix for the Number of all Types of Experiences with Robots



5.2.2.7 Relationship between participants' experience with robots and the dependant variables

Multiple Spearman rho's correlations were conducted between the dependant variables and the three types of experiences with robots to investigate if there was any relationship between them. Correlations (see Table 5.14) and visual exploration of scatterplot graphs indicated that there was a weak relationship between experience with fictional robots and the second and third subscale of the NARS only; as well as a very weak correlation between the second NARS subscale and indirect experience with non-fictional robots. Initially, it was planned to enter the three different experiences with robots as covariates into a MANCOVA as experience is often cited to have an effect on people's attitudes toward robots. However, these results suggest that this is not the case for this study and experience is unlikely to contribute to the explanatory power of the planned analysis. As such, the different types of experiences were not considered further.

Table 5.16

Correlations (Spearman rho) Between the Dependant Variables and Potential Covariates

Variables	Experience with fictional robots	Indirect experience with non-fictional robots	Direct experience with non-fictional robots
Anxiety (ANX)	.10	.05	.09
Attitude toward the use of robots (ATT)	.14	.11	-.03
Perceived enjoyment (PENJ)	.05	.00	-.08
Implicit attitudes (IAT)	-.13	.00	-.13
Interaction with robots (NARS-S1)	-.09	-.07	-.05
Social influence of robots (NARS-S2)	-.24*	-.19*	-.08
Emotion in interaction with robots (NARS-S3)	-.23*	-.10	-.08

*$p \leq .05$

5.2.2.8 Relationship between participants' experience with robots and their reported belief in the fictionality of the robots and video

Although not initially part of the planned analyses, it was decided that given the percentage of participants who did not think that the robots and video were fictional or non-fictional in accordance with the condition they were assigned to, the role of robot-related experiences on participants' belief should be explored.

Participants were asked whether they believed the story and robots depicted in the video were fictional or non-fictional. Their responses were then recoded as either "matches condition" or "does not match condition". For example, participants in the fictional condition who reported that they believed the robots in the video to be fictional were assigned a values of 1 that corresponded to the label "matches condition". A Mann-Whitney U test was conducted for each of the three types of experience with robots and the two dichotomous variables: belief in the fictionality of the story and belief in the fictional status of the robots.

A Bonferonni correction was applied to account for the multiple tests and a new critical *p* value of .017 (3 comparisons per question) was used. Distribution of the number of experiences with fictional robots was not statistically different between participants who believed the robots to be fictional/non-fictional in accordance with their condition assignment and participants who did not, $U = 1737.5$, $z = -2.18$, $p = .029$. The same was true for indirect experiences with non-fictional robots, $U = 2047$, $z = -0.80$, $p = .422$; and for direct experiences with non-fictional robots, $U = 1841.5$, $z = -1.83$, $p = .067$.

Distribution of the number of experiences with fictional robots was not statistically different between participants who believed the video to be fictional/non-fictional in accordance with their condition assignment and participants who did not, $U = 1963$, $z = -1.42$, $p = .155$. The same was true for indirect experiences with non-fictional robots, $U = 2080$, $z = -0.911$, $p = .363$; and for direct experiences with non-fictional robots, $U = 1974$, $z = -1.47$, $p = .143$.

5.2.3 Discussion

The first question this study aimed to answer was whether it was possible to manipulate the extent to which participants believed that robots depicted in visual media (i.e., a video) are fictional (i.e., made-up) or non-fictional (i.e., real). Following multiple pilot studies described in this chapter, a method capable of affecting people's beliefs about the fictional status of robots was developed and tested. The findings from the present study confirm that participants' beliefs regarding robots are to some extent malleable and the potential impact of this finding will be discussed in the next section. The second question this study, and indeed this chapter, aimed to answer was whether people's perception of the fictionality of robots has an effect on their attitudes toward robots. Findings suggest that this is likely the case and this too will be discussed in detail further down in this section.

5.2.3.1 Malleability of participants' belief in the fictionality of robots

Approximately half or more of the participants believed that video that they watched and the robots depicted within were fictional or non-fictional in accordance with the condition to which they were assigned. This suggests that it is possible to convince some people that previously unseen robots are either a fictional creation or technology that exists in real-life by providing them with brief descriptive text prior to viewing said robots. Although no specific hypotheses were set regarding the malleability of participants' belief in the fictionality of the robots, there was an expectation that such a manipulation would be possible with a non-specialist sample (i.e., people with no professional background in robotics). This was predominantly driven by the idea that the majority of the general public is unlikely to have had much experience or contact with real robots but could have instead been exposed to various representations of robots in popular media such as films.

As discussed in the Introduction of this chapter (see Section 5.1), representations of robots in films (and likely other media as well) are rarely reflective of the abilities of currently available robots (Kriz et al., 2010). Therefore, it was expected that a naïve participant's lack of knowledge about the abilities and use of robotics would result in malleable beliefs about the fictionality of unfamiliar robots. As established in the Results section, participants in this study reported, on average, a low number of experiences with fictional robots and a low number of indirect and direct experiences with non-fictional robots. Interestingly, however, the number of any type of experience with robots had no significant relationship with whether participants believed the condition dependent information they were given. This puts some doubt over the initial assertion that prior experience or knowledge about robots (non-fictional robots in particular) would result in less malleable beliefs about the fictional status of unfamiliar robots. It should also be noted that the vast majority of the participants had no direct experience with non-fictional robots, little indirect experience, and a somewhat more varied experience with fictional robots. In fact, less than 5% of the participants had a large number of experiences that would have implied a professional background in robotics or an exceptionally high number of direct contact with robots. As such, it is likely that this study's sample lacks the sufficient size and diversity, and thus power, to detect any relationship between prior experience with robots and participant's belief in whether the robots were fictional or not. While the effect of prior experience with non-fictional robots on people's attitudes has been investigated before, it remains unclear how such experiences affect people's knowledge about robotics.

There is an underlying that fictional representations of robots are typically unrealistic and contribute to the formation of attitudes that are not grounded in the reality of robotics as it is currently. However, there is no evidence that direct and indirect experiences with robots are inherently representative of the true abilities and functions of robotic systems. As of yet, robots are not encountered frequently by the general population in most countries. That may mean that people who have experienced direct contact with robots might have done so as participants in research or at some sort of public engagement event. This can be somewhat problematic as some HRI studies purposely deceive participants as to the true abilities of robots (e.g., more autonomy than it is currently possible; Gaudiello et al., 2016) or, unknowingly, present robots in a not entirely representative of reality way (e.g., Conti et al., 2017). Unfortunately, most studies do not report their debriefing content and process in enough detail for a judgement to be made regarding the extent to which researchers rectify intentional and unintentional misrepresentations of robots' abilities and purpose. Public engagement events can be equally wroth with unintentional and intentional misrepresentation of robots' true abilities. The purpose of such events is to engage with the general public, usually via some entertaining demonstration of the robots' abilities. A mixture of pre-programmed responses and behind-the-scenes teleoperation, plus little time left to educate the public about the current state of robotics and the true abilities of robots, could be contributing to a somewhat unrealistic representation of such technology. Since studies generally do not formally ask participants to report on their knowledge or assumptions about robots following interactions and neither do exhibitioners at public engagement events, a formal investigation of the way such interactions are interpreted by the general public is needed. Media coverage of robots and advances in robotics can also be problematic as it can range from incredibly optimistic (e.g., robots as domestic helper in the very near future) to downright paranoid in nature (e.g., robots taking human jobs; Pettit, 2018). Furthermore, Bruckenberger et al. (2013) found that the majority of their participants were not familiar with robots represented in mass media. Those who had encountered media coverage of non-fictional robots did not generally accept them as prototypical robots. Although this thesis focuses on the effect of fictional representations of robots on people's attitudes and beliefs, the effect of non-fictional representations should also be explored further.

While there may be little practical use of manipulating people's beliefs about the fictionality of specific robots, findings from this study tell us something about the stability of people's beliefs about robotics and the general public's knowledge about fictional and non-fictional robots. It may also serve as a call for more transparency in HRI research and non-fictional media in order to avoid overly optimistic or pessimistic beliefs about robotics.

5.2.3.2 Effect of perceived fictionality on explicit attitudes toward the robots in the video

It was found that, in general, participants who were told that the video and robots were non-fictional had more positive attitudes toward the robots in the video than participants who were told the video and robots were fictional. Findings for each subscale are discussed in detail below.

Attitude (ATT subscale). Although not explicitly stated by the developers of this scale (Heerink et al., 2010), the attitude subscale can be considered as a measure of participants' cognitive attitudes since the scale items are about participants' thoughts on whether the robots should be used. On average, participants reported positive attitudes toward the use of the robots in the video, regardless of the condition to which they were assigned. However, participants who were assigned to the non-fictional condition reported significantly more positive attitudes toward the robots than did participants in the fictional condition. This medium sized effect of condition assignment suggests that believing robots to be non-fictional leads to more positive attitudes toward their use than believing the same robots to be fictional. Results from this study are somewhat in line with findings by Mubin et al. (2015) who observed that YouTube videos of two non-fictional robots (Nao and Shakey) generated more engagement and positive interest (i.e., positive comments) compared to the videos of the fictional robots (HAL900 and Astro Boy). This may be for a variety of reasons. It could be that fictional characters (including robots) may be more prone to negative comments due to their affiliation with particular sci-fi movies and their role within them. Alternatively, it could be that nonfictional robots are less familiar to viewers and, as such, are more novel, thus leading to more engagement with videos depicting non-fictional robots. Although these explanations may be applicable to Mubin et al.'s study where the fictionality of the robots was not manipulated, it does not necessarily explain why identical and unfamiliar robots as the ones in the present study would be subject to such biases. Since this particular scale deals with attitudes toward the use of specific robots, one plausible explanation could be that participants found it easier to imagine using the robots in the video that were labelled as non-fictional and thus had more positive attitudes toward their use. This, however, is impossible to confirm without collecting additional data.

Anxiety (ANX subscale). After being asked to imagine interacting with the robots in the video, participants, on average, reported that they would not feel anxious during such an interaction. This is not a particularly surprising finding as existing research, on average, supports neutral levels of anxiety toward interaction with robots during indirect HRI (see Chapter 2). This level of neutrality was not significantly different between the two conditions, suggesting that belief in the fictionality of the robots and video had no effect on participants' anxiety toward a hypothetical interaction with the robots. This could very well be because while the information regarding the robots was different, all participants watched the same scenario which could very well have resulted in participants imagining similar low-anxiety interactions with the robots. Since this was not an imagined contact exercise and we did not collect data on what participants' imagined (if anything), it is difficult to interpret these findings. Surprisingly, anxiety was also not correlated with any type of experience with robots. Existing research has shown that prior experience with non-fictional robots leads to reduced anxiety during direct HRI (e.g., Bartneck et al., 2007). Of course, since participants did not interact with the robots in the video directly, relying on existing research to interpret these findings is not sufficient.

Perceived enjoyment (PENJ subscale). After being asked to imagine interacting with the robots in the video, participants, on average, reported that they would somewhat enjoy such an interaction. Similar to anxiety, the levels of perceived enjoyment across the two conditions was the same, suggesting that whether participants believed the robots to be fictional or not had no effect on their enjoyment in an imaginary scenario. Although not made explicitly clear by the developers of this scale (Heerink et al., 2010), perceived enjoyment could be considered as a measure of participants' affective attitudes since the scale items are about emotions toward the robots. This would in turn suggest that while the perceived fictionality of robots affects people's cognitive attitudes, it does not affect their affective attitudes toward the same robots.

5.2.3.3 Effect of perceived fictionality on attitudes toward the robots in general

It was found that, in general, participants who were told that the video and robots were non-fictional had more positive attitudes toward robots in general than participants who were told the video and robots were fictional. Findings for each measure are discussed in detail below.

Interaction with robots (NARS-S1). Participants in both conditions reported, on average, slightly positive attitudes toward robots which is consistent with findings from the broader literature (see Chapter 2). This particular scale measures people's negative attitudes toward interacting with robots in general and as such is similar to the attitude (ATT) measure. Results were also similar as participants in the non-fictional condition reported significantly more positive attitudes than participants in the fictional condition. This implies that believing robots to be non-fictional leads to more positive attitudes, not only toward specific robots, but also toward robots in general. As described in the previous section, these findings are somewhat consistent with results from Mubin et al. However, they did not make any specific observations about attitudes toward robots in general so this connection is tenuous at best. One possible explanation could be that, for whatever reason, watching what participants believed to be non-fictional robots elicited more positive associations with robots in general than did watching what participants believed to be fictional robots. Although participants were asked if they had anything in mind while answering the NARS, the vast majority reported that they did not think or imagine anything while answering the questions. From those who answered the question, there appeared to be no difference between the conditions in the number of fictional or non-fictional robots participants thought about or in the number of positive, negative or neutral comments they left. However, it is difficult to make any conclusion based on this limited data set. Further investigation of what people think about when answering questions about robots is needed to inform this finding.

Another similarity between NARS-S1 measure and the attitude (ATT) scale is that there was no apparent relationship between attitudes toward interaction with robots in general and the number of experiences participants had with fictional and non-fictional robots. This is inconsistent with findings by Riek et al. (2011) who noted a weak negative correlation between number of fictional films seen by participants and the NARS, indicating that participants who watched more films which depicted robots were also more likely to have more positive attitudes toward robots in general. Riek et al. did not however differentiate between the different subscales of the NARS so it is entirely possible that there was no correlation between attitudes toward interaction with robots and experience with fictional robots. Additionally, as mentioned previously the data we collected on participants' number of robot-related experiences was extremely positively skewed and leptokurtic which makes it difficult to interpret. A separate investigation with a considerably larger and more diverse sample is likely needed to understand the impact of people's robot-related experiences on their attitudes toward robots.

Social influence of robots (NARS-S2). Participants in the non-fictional conditions, on average, reported nearly neutral attitudes toward the social influence of robots while participants in the fictional condition reported slightly negative attitudes toward robots. While this pattern is consistent with participants' self-reported attitudes so far (ATT and NARS-S1 subscales), no statistical difference between the two conditions was found in this instance. This implies that participants' belief in the fictionality of the robots had no effect on their attitudes toward the broader social impact of robotics.

This particular scale was one of two scales which were weakly negatively correlated with the number of fictional experiences participants reported. This suggests that the more experience people have with fictional robots the more positive their attitudes toward the social influence of robots. This is consistent with Riek et al. (2011) who found similar correlation between the NARS and number of fictional films viewed by participants. It is entirely possible that people who tend to watch films that contain robots are also interested in robotics and already have more positive attitudes toward robots and their social influence. This idea is partially supported by the fact that we found a moderate positive correlation between the number of experiences with fictional robots and the number of indirect experiences with non-fictional robots. No such relationship was found with direct experience of non-fictional robots but this may be due to our sample (i.e., predominantly lay persons) and the relative rarity of direct interaction with robots. Of course, there is only so much that can be inferred from a correlational analyses and more research is needed to understand the relationship between different robot-related experiences and attitudes toward robots.

Emotion in interaction with robots (NARS-S3). Although not specifically stated, the third subscale of the NARS essentially measure people's affective attitudes toward robots in general as the scale items refer to the feeling participants think they will experience when interacting with robots. Participants in the non-fictional condition on average reported slightly positive attitudes toward emotion when interacting with robots while participants in the fictional condition reported slightly negative attitudes toward robots. This difference was significant which is somewhat surprising given that no such difference was found between conditions for the perceived enjoyment (PENJ) scale which measured participants' affective attitudes toward the robots in the video. Similar to the NARS-S1, one possible explanation could be that, for whatever reason, watching what participants believed to be non-fictional robots elicited more positive associations with robots in general than did watching what participants believed to be fictional robots which in turn could have resulted in more positive attitudes for the non-fictional group. However, as stated before, we found no evidence of this based on the answers participants provided to the qualitative question following the NARS.

Similar to the NARS-S2 subscale, there was a weak negative correlation between participants' emotion in interacting with robots and the average number of experiences with fictional robots they reported. This suggests that the more experience someone has with fictional representations of robots, the more positive attitudes they are likely to have in relation to their expected emotions during interaction with robots in general. However, these findings are, as mentioned previously, difficult to interpret.

Implicit attitudes/associations. In general, participants made quicker associations between the human images and pleasant words, and robot images and unpleasant words. This was only slightly so, meaning that participants did not have particularly strong negative associations with robots and preferred humans only slightly. These findings are difficult to interpret given the relative rarity of implicit measures in the robotics literature although not necessarily surprising as this particular study did not aim to influence people's attitudes toward robots relative to their attitudes of humans. In addition, implicit attitudes were not correlated with any of the other measures. This is at odds with findings from MacDorman, Vasudevan, and Ho (2008) who observed a weak positive correlation between d-scores and self-reported measures of attitudes (i.e., NARS). Although, the average d-scores they found are very similar to the ones found in this study (i.e., slight preference for humans over robots).

5.2.3.4 Limitations and future directions

One major limitation of the present study is the lack of a control group. We do not know whether participants would assume the robots to be fictional or non-fictional when not provided with any information pertaining to their fictionality. This could be problematic as the video and robots may be more prototypical robot representations of non-fictional rather than fictional robots. In other words, participants in the fictional condition may have had less positive attitudes toward robots because what they observed was not a 'typical' robot that may appear in fiction. As described in the results, participants, on average, reported more experiences with fictional robots than with non-fictional robots, meaning that they were less likely to have in mind a prototypical model of a real robot to which they could compare the ORO, DORO, and CORO robots. Since the ‘typicality’ of the robots in the video was not evaluated, another study is needed to inform this conclusion. Additionally, according to Mubin et al. (2015), people prefer less human-like features for non-fictional robots and more human-like features for fictional robots (i.e., people like more humanoid fictional robots and humanoid non-fictional robots). It could, therefore, be possible that the design of the ORO, DORO, and CORO robots might have affected participants’ attitudes toward the robots in the video depending on the condition. This may have contributed to participants having more positive attitudes toward the robots they believed to be non-fictional since the design of the robots (especially the level of anthropomorphism) is arguably more in line with real rather than fictional robots. Lastly, it is clear that more research is needed to inform the relationship between the experiences people have with fictional and non-fictional robots and their effect on people’s attitudes (see Chapter 6).

5.3 Pilot Study 4 - Selection of Images of Fictional and Non-Fictional Robots

This pilot study's primary aim was to select five (or more) fictional and non-fictional robot images for use in Study 6 which sought to prime participant with images of fictional and non-fictional robots prior to measuring their explicit and implicit attitudes. The aim was to match the fictional and non-fictional images of robots on six dimensions (*pleasant* - *unpleasant, friendly* - *unfriendly, safe* - *threatening, familiar* - *unfamiliar, real* - *fictional, and good* - *evil*) as rated by a student sample. It was expected that images of fictional robots will be rated as more fictional than images of non-fictional robots. No other predictions were made.

5.3.1 Method

5.3.1.1 Participants

A sample of 159 participants completed the study. The majority of participants (N = 144) were recruited via a mailing list of student volunteers at the University of Sheffield. As a compensation for their time, participants were entered into a draw for a £25 Amazon voucher which was randomly awarded after the end of the data collection period. First year psychology undergraduates (N = 15) were also recruited via the Department of Psychology's Online Research Participation System (ORPS). These students received two course credit as compensation for their time.

5.3.1.2 Materials

Fourteen fictional humanoid robots were selected from a list of top-ranking movies compiled by Sandoval et al. (2014). Since other movies containing robot characters have been created since Sandoval et al.'s study was published, an online search of movies was conducted and the top ten highest ranking movies were also considered. From these movies, seven humanoid robots which were represented negatively (i.e., characters which work directly against the main character(s) of the story, usually in a violent or deceitful way) and seven robots which were represented in a positive way (i.e., characters which were helpful to the main character of the story) were selected. An effort was made to select humanoid robots with similar features (e.g., head, neck, shoulders, etc.). The fourteen non-fictional humanoid robots were selected by searching for robot related news articles via Google News and selecting robots from the most recent articles.

All images were edited so that the full body of the robots was visible on a white background. The full list of robots can be found in Appendix O.

5.3.1.3 Procedure

Data was collected online via Qualtrics. All participants were asked to rate the 28 images of fictional and non-fictional robots on six seven-point bi-polar scales (pleasant-unpleasant, friendly-unfriendly, safe-threatening, familiar-unfamiliar, real-fictional, good-evil, and humanlike-mechanical). Images appeared individually and in random order with the six scales presented in random order directly below the image. Participants were told that there were no right or wrong answers. After rating all of the images, participants were debriefed (see Appendix D for debrief procedure).

5.3.2 Results

Details about exclusions, treatment of the data, and assumption checks relevant to the analyses carried out in this section can be found in Appendix P.

5.3.2.1 Relationship between the variables

The Pleasant-Unpleasant, Hostile-Friendly, Safe-Threatening, and Evil-Good dimensions were all strongly correlated (although less so for the non-fictional robots; see Table 5.17) and fictional robots were, on average, rated as slightly more unpleasant (M = 4.04, SD = 1.59; $t(158)$ = 9.81, p < .001), less friendly (M = 3.96, SD = 1.85; $t(158)$ = -16.28, p < .001), more threatening (M = 4.09, SD = 1.81; $t(158)$ = 19.06, p < .001), and less good (M = 4.09, SD = 1.80; $t(158)$ = -12.71, p < .001) than the non-fictional robots (M = 3.48, SD = 0.61; M = 4.83, SD = 0.59; M = 3.05, SD = 0.57; M = 4.83, SD = 0.48, respectively).

The Unfamiliar-Familiar dimension was only weakly to moderately correlated (Table 5.17) with the other variables which could be the result of differences between the mean ratings of the two groups of robots. In general, participants rated the fictional robots as more familiar (M = 4.81, SD = 1.03) than the non-fictional ones (M = 3.36, SD = 0.72), $t(158)$ = 13.98, p < .001.

The Real-Fictional dimension correlated very weakly (Table 5.17) with the other variables which is not too surprising given that, on average, participants rated the fictional robots as more fictional (M = 6.18, SD = 0.60) than the non-fictional robots (M = 2.80, *SD* = 0.78), *t*(158) = 32.88, *p* < .001. This dimension did however correlate moderately with the Unfamiliar-Familiar and Evil-Good dimensions which could be a result of fictional robots being rated as both more familiar and more evil than non-fictional ones.

The Humanlike-Mechanical dimension correlated very weakly (Table 5.17) with all variables and more so for the fictional robots (*M* = 4.25, *SD* = 0.95) which, on average, were not rated differently than the non-fictional robots (*M* = 4.36, *SD* = 0.89), *t*(158) = -1.88, *p* = .062. All *t*-test were conducted with a Bonferroni adjusted critical alpha level of .007 (.05/7) to account for multiple comparisons.

Table 5.17

Correlations (Pearson's r) Between All Variables

Variables	P-U	H-F	S-T	U-F	R-F	E-G	H-M
Pleasant-Unpleasant	-	-	-	-	-	-	-
Hostile-Friendly							
Overall	-.82*	-	-	-	-	-	-
Fictional	-.87*						
Non-fictional	-.67*						
Safe-Threatening							
Overall	.81*	-.87*	-	-	-	-	-
Fictional	.88*	-.93*					
Non-fictional	.66*	-.67*					
Unfamiliar-Familiar							
Overall	-.36*	.32*	-.27*	-	-	-	-
Fictional	-.48*	.50*	-.47*				
Non-fictional	-.39*	.33*	-.33*				
Real-Fictional							
Overall	.11*	-.14*	.23*	.36*	-	-	-
Fictional	-.12*	.13*	-.10*	.33*			
Non-fictional	.11*	-.09*	.19*	-.09*			
Evil-Good							
Overall	-.80*	.89*	-.84*	.34*	-.13*	-	-
Fictional	-.86*	.93*	-.90*	.52*	.14*		
Non-fictional	-.64*	.76*	-.67*	.32*	-.13*		
Humanlike-Mechanical							
Overall	.16*	-.13*	.11*	-.19*	-.09*	-.13*	-
Fictional	.12*	-.11*	.12*	-.09*	.012	-.12*	
Non-fictional	.29*	-.26*	.22*	-.26*	-.01	-.22*	

Note. P-U denotes the pleasant-unpleasant dimension; H-F denotes the hostile friendly dimension; S-T denotes the safe-threatening dimension; U-F denotes the unfamiliar-familiar dimension; E-G denotes the evil-good dimension; H-M denotes the humanlike-mechanical dimension.

*correlations significant at $p < .01$

5.3.2.2 Selection of the robot images

The means and standard deviations (see Table 5.18) of the ratings participants gave each robot were calculated and robots with mean ratings between 3 and 5 were excluded (three fictional – Optimus Prime, and T-X, Marvin; four non-fictional – NAO, Romeo, Kengoro, Valkyrie, and Zeno) as they were, on average, rated as somewhat more ambiguous in their fictionality and were considered not to be appropriate for the aims of the study (i.e., selecting five fictional and five non-fictional robot images to use for a priming task) and as such discounted. There was a significant difference of the real-fictional ratings between the 11 fictional and 9 non-fictional robots, as indicated by a significant paired-samples t-test, $t(158) = -39.3$, $p < .001$. It was also noted that, on average, none of the fictional robots were rated as highly unpleasant, threatening, hostile, or evil (indicated by mean ratings between 1 and 3); whereas, three of the fictional robots (Megatron, T-800, and Ultron) were all rated as highly unpleasant, threatening, hostile, and evil. As such, it was decided not to include these three robots in any further analysis as there were no non-fictional robots that were rated in a similar way.

Table 5.18

Descriptive Statistics for all 28 images for the Pleasant – Unpleasant, Hostile – Friendly, Safe – Threatening, Unfamiliar – Familiar, Real – Fictional, Evil – Good, and Humanlike – Mechanical dimensions (1-7)

		Pleasant – Unpleasant		Hostile – Friendly		Safe – Threatening		Unfamiliar – Familiar	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Fictional Robots	Baymax	1.79	1.22	6.40	1.00	1.66	1.13	5.91	1.64
	BumbleBee	3.67	1.83	4.47	1.85	3.92	1.91	5.17	2.00
	C3P0	2.67	1.65	5.96	1.36	2.19	1.42	5.97	1.69
	Cyberman	5.41	1.44	2.50	1.54	5.40	1.54	5.22	2.03
	Marvin	3.00	1.44	4.96	1.40	2.75	1.36	4.12	2.12
	Megatron	6.28	1.13	1.40	0.85	6.64	0.86	4.05	2.30
	Optimus Prime	4.41	1.56	3.57	1.31	4.43	1.45	4.18	2.05
	Robocop	4.00	1.82	4.24	1.86	4.22	1.84	5.30	1.95
	Rodney	4.22	1.46	3.78	1.47	4.16	1.60	4.72	2.04
	Sonny	2.05	1.21	6.24	1.06	1.92	1.15	5.64	1.71
	T-800	6.06	1.19	1.54	0.84	6.42	0.91	4.31	2.21
	T-X	5.22	1.39	2.40	1.14	5.54	1.14	2.68	1.57
	Ultron	5.92	1.32	1.67	0.90	6.35	0.90	3.73	2.24
	WALL-E	1.86	1.27	6.37	1.08	1.69	1.06	6.39	1.20
Non-fictional Robots	ARMAR-6	3.41	1.28	4.68	1.15	3.10	1.44	3.18	1.62
	Asimo	2.61	1.21	5.47	1.24	2.28	1.21	5.22	1.85
	Atlas	4.19	1.49	3.99	1.38	3.79	1.53	3.17	2.08
	HUBO	3.96	1.27	4.31	1.18	3.12	1.32	2.62	1.55
	Justin	3.63	1.44	4.72	1.28	3.32	1.27	2.77	1.44
	Kengoro	4.52	1.40	3.87	1.34	4.14	1.47	2.62	1.55
	NAO	2.46	1.25	5.81	1.04	2.17	1.20	4.12	2.05
	Pepper	2.65	1.31	5.62	1.10	2.24	1.09	4.15	1.79
	Robina	3.73	1.53	4.71	1.27	3.10	1.32	2.82	1.59
	Robothespian	3.92	1.45	4.75	1.24	3.33	1.34	2.91	1.59
	Romeo	3.09	1.63	5.34	1.24	2.72	1.33	3.37	1.81
	THR3	3.47	1.37	4.51	1.32	3.22	1.42	3.58	1.70
	Valkyrie	3.63	1.36	4.66	1.19	3.40	1.26	3.20	1.61
	Zeno	3.45	1.77	5.22	1.36	2.82	1.49	3.24	2.01

Table 5.18 (continued)

		Real – Fictional		Evil – Good		Humanlike – Mechanical	
		M	*SD*	*M*	*SD*	*M*	*SD*
Fictional Robots	Baymax	6.55	1.12	6.41	1.01	3.65	1.57
	BumbleBee	6.68	0.97	4.62	1.82	5.38	1.60
	C3P0	6.46	1.29	5.72	1.51	3.72	1.79
	Cyberman	6.06	1.50	2.63	1.48	4.59	1.67
	Marvin	5.46	1.82	5.07	1.24	4.15	1.60
	Megatron	6.62	1.03	1.47	0.88	5.56	1.63
	Optimus Prime	4.90	2.09	3.89	1.29	3.39	1.66
	Robocop	6.67	0.94	4.54	1.89	5.22	1.66
	Rodney	6.21	1.34	4.30	1.50	2.72	1.61
	Sonny	6.75	0.63	6.25	1.02	3.26	1.73
	T-800	6.13	1.55	1.69	1.12	4.31	1.95
	T-X	5.15	1.70	2.58	1.16	3.85	1.81
	Ultron	6.32	1.21	1.74	1.07	3.89	1.88
	WALL-E	6.55	1.32	6.38	1.05	5.87	1.52
Non-fictional Robots	ARMAR-6	2.18	1.45	4.83	1.23	5.38	1.36
	Asimo	1.78	1.21	5.46	1.21	3.78	1.70
	Atlas	2.11	1.43	4.13	1.32	4.13	1.32
	HUBO	2.16	1.38	4.57	1.11	5.62	1.34
	Justin	2.89	1.45	4.77	1.17	5.30	1.34
	Kengoro	3.45	1.66	3.96	1.33	4.03	1.63
	NAO	3.03	1.86	5.62	1.16	4.02	1.61
	Pepper	2.23	1.41	5.43	1.20	4.26	1.54
	Robina	2.96	1.48	4.73	1.25	4.72	1.50
	Robothespian	2.75	1.54	4.57	1.26	4.30	1.60
	Romeo	3.13	1.60	5.23	1.33	3.09	1.62
	THR3	2.16	1.28	4.55	1.26	4.43	1.59
	Valkyrie	3.73	1.70	4.70	1.14	3.69	1.36
	Zeno	4.61	2.08	5.01	1.36	2.81	1.70

Analysis and image selection. Given that some variables were severely skewed and kurtotic (see Appendix Q) and that normalising the distribution of all variables is unlikely to be achieved via transformation of the data, and that we have decided to retain some unusual cases (i.e., data points), it was decided that a nonparametric test would be more appropriate for this particular data set.

Multiple Friedman's ANOVA tests were conducted to investigate which of the ten robots were rated similarly by participants on the six dimensions (pleasant-unpleasant, hostile-friendly, safe-threatening, familiar-unfamiliar, evil-good, humanlike-mechanical). It was found that there was a significant difference between the ratings given to the robots on all six dimensions (see Table 5.19).

Table 5.19

Friedman's ANOVA Statistics for Each Dimension (Fictional vs. Non-Fictional Robots)

	N	χ^2	df	p
Pleasant - Unpleasant	139	833.67	16	.001
Hostile - Friendly	139	977.72	16	.001
Safe - Threatening	134	880.94	16	.001
Unfamiliar - Familiar	142	921.52	16	.001
Evil - Good	140	864.75	16	.001
Humanlike - Mechanical	142	820.05	16	.001

Step-down analyses were conducted to follow-up each significant Friedman's test in order to detect which robots were not significantly different from each other. Robots that were rated in a similar way by participants were grouped together and each homogeneous group (all non-significant, $p \geq .067$) were explored to see whether they contained a mix of fictional and non-fictional robots. Where this was the case, the fictional and non-fictional robot(s) and the dimension on which they matched (i.e., were not significantly different from each other) were noted down in a matrix (see Table 5.20). Since we were not interested in whether robots from the same group (i.e., fictional or non-fictional) were significantly different from each other, this information was not noted down or explored further.

In order to select five fictional and five non-fictional robots that were rated as similarly as possible by participants, the fictional - non-fictional robot pairings that matched on the most number of dimensions (e.g., Justin and Bumblebee; see Table 5.20) were compared using their median ratings for each dimension (see Table 5.21). Ideally, five different fictional robots would have been matched with five different non-fictional robots. However, the fictional robots Baymax, Cyberman, and Wall-E appeared to be quite different in terms of the way they were rated compared to all of the non-fictional robots (see Table 5.20). Since this was the case, it was decided that they would not be used for the priming task. This meant that it was not possible to have individual matching pairs of images for all of the selected robots. In other words, the non-fictional robot Asimo was matched with not only Sonny but also with C3P0. Similarly, the fictional robot Bumblebee was matched with two non-fictional robots, Justin and Robothespian. Robocop and ARMAR-6 were selected as a matching fictional – non-fictional robot pair as were Rodney and Atlas.

Table 5.20

Dimensions on Which the Fictional (First Row) and Non-Fictional (First Column) Robots Match

	Baymax	Bumblebee	C3P0	Cyberman	Robocop	Rodney	Sonny	WALL-E
ARMAR-6	-	P-U, H-F, E-G, H-M	-	-	P-U, H-F, E-G, H-M	E-G	-	-
Asimo	H-M	U-F	P-U, S-T, E-G, H-M	U-F	U-F	U-F	S-T, U-F, H-M	-
Atlas	-	S-T, H-M	-	-	P-U, H-F	P-U, H-F, E-G	-	H-M
HUBO	-	H-F, E-G, H-M	-	-	P-U, H-F, E-G	P-U, E-G	-	H-M
Justin	-	P-U, H-F, S-T, E-G, H-M	-	-	P-U, H-F, E-G, H-M	-	-	-
Pepper	-	-	P-U, S-T, E-G, H-M	H-M	-	U-F	S-T	-
Robina	-	P-U, H-F, E-G	-	H-M	P-U, H-F, E-G	-	-	-
Robothespian	-	P-U, H-F, S-T, E-G	-	H-M	P-U, H-F, E-G	P-U, E-G	-	-
THR3	-	P-U, H-F, S-T, E-G	-	H-M	H-F, E-G	E-G	-	-

Note. Grey boxes signify robots that were rated similarly by participants and as a result were selected for the priming task. P-U denotes the pleasant-unpleasant dimension; H-F denotes the hostile friendly dimension; S-T denotes the safe-threatening dimension; U-F denotes the unfamiliar-familiar dimension; E-G denotes the evil-good dimension; H-M denotes the humanlike-mechanical dimension.

Table 5.21

Median Rating of Each Robot for Each of the Six Dimensions

		Pleasant – Unpleasant	Hostile – Friendly	Safe – Threatening
Fictional Robots	Baymax	1	7	1
	BumbleBee	3	5	4
	C3P0	2	7	2
	Cyberman	5	2	6
	Robocop	4	4	5
	Rodney	4	4	4
	Sonny	2	7	2
	WALL-E	1	7	1
Non-fictional Robots	ARMAR-6	3	5	3
	Asimo	2	6	2
	Atlas	4	4	4
	HUBO	4	4	3
	Justin	3	5	3
	Pepper	3	6	2
	Robina	4	5	3
	Robothespian	4	5	3
	THR3	3	4	3
		Unfamiliar – Familiar	Evil – Good	Humanlike – Mechanical
Fictional Robots	Baymax	7	7	4
	BumbleBee	6	5	6
	C3P0	7	6	3
	Cyberman	6	2	5
	Robocop	6	5	6
	Rodney	5	4	2
	Sonny	6	7	3
	WALL-E	7	7	6
Non-fictional Robots	ARMAR-6	3	5	6
	Asimo	6	6	3
	Atlas	2	4	6
	HUBO	2	4	6
	Justin	2	5	6
	Pepper	4	5.5	4
	Robina	2	5	5
	Robothespian	3	5	4
	THR3	3	4	4

5.3.3 Discussion

The aim of this study was to select five fictional and five non-fictional robot images to be used as the priming stimuli in the proposed experimental study. Since a between-subjects design with two conditions (priming with fictional robot images vs. priming with non-fictional robot images) was to be used in Study 6, it was important to match the images on a number of dimensions (*pleasant - unpleasant, friendly - unfriendly, safe - threatening, familiar - unfamiliar, good - evil, and humanlike - mechanical*) in order to limit the effect of those particular robot characteristics on the dependant variables. While the average ratings given to the robot images by participants did not allow for a complete matching between the fictional and non-fictional robots, the closest possible match was found using the method described in the previous section. As a result, five fictional (BumbleBee, C3P0, Robocop, Rodney, and Sonny) and five non-fictional (ARMAR-6, Asimo, Atlas, Justin, and Robothespian) robots were selected. Their 'fictionality' was confirmed by participants' average ratings on the fictional-real dimension, with fictional robots being rated consistently low on this dimension (i.e., "from a film or TV programme") and non-fictional robots being rated consistently high (i.e., "a real piece of technology").

It was impossible to match fictional and non-fictional robots closely on the unfamiliar-familiar dimension and the majority of other dimensions. This was not particularly surprising given that, on average, participants rated the fictional robots as significantly more familiar than the non-fictional robots. This was most likely a result of fictional robots being much more accessible through popular media (i.e., films) than non-fictional robots. Since familiarity (or experience) with robots has been suggested as a possible factor in predicting people's attitudes and behaviour toward robots (Nomura et al., 2006; MacDorman et al., 2009; Stafford et al., 2010) this was of some concern for Study 6. It was therefore necessary to measure each participants' familiarity of the robots they see in the priming task and consider familiarity as a moderating variable in Study 6.

It was also not possible to match robots closely on the humanlike-mechanical dimension. Although there appeared to be no significant difference between the ratings participants gave the fictional and non-fictional robots, and the humanlike-mechanical dimension correlated poorly with other factors, this mismatch was still not ideal. Studies have shown that the appearance of robots can have an effect on people's attitudes especially where humanoid robots are concerned (Fink, 2012; Li et al., 2010). As such, it was necessary to consider human-likeness as a possible moderating variable in Study 6.

In conclusion, the final selection of fictional and non-fictional robot images can be considered similar enough (as rated by participants) to be used in the proposed priming task. However, the results from this study also suggested that fictional robots consistently differ from non-fictional robots on some dimensions (e.g., *familiarity*) which may affect people's attitudes toward those robots. This difference was therefore taken into consideration during Study 6 by asking participants to rate the robots using the dimensions from this pilot study.

5.4 Study 6 - The Effect of Priming Participants with Fictional and Non-Fictional Robots on Their Implicit Attitudes Toward Robots

A somewhat recent study by Thellman and Ziemke (2017) suggests that measuring people's attitudes toward robots may be affected by the examples of robots that people have in mind when answering such questions. They randomly assigned participants to complete identical questionnaires which depicted one of three possible robots at the top of each questionnaire. The robots were either non-, semi-, or highly anthropomorphic (i.e., they had a human-like appearance to a lesser or greater extent). Participants' attitudes were measured using the Negative Attitudes toward Robots Scale (NARS; Nomura et al., 2004). Thellman and Ziemke found that presenting different images of robots resulted in significant differences on some dimensions of the NARS suggesting that attitudes toward robots could be manipulated by presenting people with different examples of robots (at least in the short term and as measured by the NARS).

Anthropomorphism is, however, not the only way in which representations of robots may differ and subsequently influence people's attitudes toward robots. For example, the fictional status of the robot images with which participants were primed. The present research therefore sought to replicate and expand on Thellman and Ziemke's study by presenting participants with images of fictional and non-fictional robots prior to measuring their explicit and implicit attitudes. In the present study, participants' explicit attitudes were measured via NARS and their implicit attitudes were measured using the Implicit Association Test (IAT). Examples of fictional and non-fictional robots were covertly primed in order to avoid any response bias by asking participants to indicate their preferences for the design of five fictional or non-fictional robots for use in different contexts (e.g., healthcare, education, etc.). Based on Thellman and Ziemke's findings, it was expected that there would be a difference in participants implicit and explicit attitudes depending on the images of robots they viewed (fictional or non-fictional) prior to measuring their attitudes.

5.4.1 Method

5.4.1.1 Participants

Fifty-six participants completed the study (43 females; M_{age} = 29.8, SD_{age} = 13.2).[16] Table 5.22 provides the demographic details of the sample split between the two conditions. Half of the participants (N = 28) were recruited via a mailing list of staff volunteers at the University of Sheffield. As a compensation for their time, participants were entered into a draw for a £25 Amazon voucher which was randomly awarded after the end of the data collection period. First year psychology undergraduates (N = 28) were also recruited via the Department of Psychology's Online Research Participation System (ORPS). These students received three course credit as compensation for their time.

Table 5.22

Participants' Demographic Characteristics

Characteristic	Fictional condition (N = 28)	Non-fictional condition (N = 28)
Age		
M (SD)	29.29 (13.37)	30.32 (13.35)
Range	18 - 63	18 - 60
Gender Identity		
Female (%)	20 (71.4)	23 (82.1)
Male (%)	8 (28.6)	4 (14.3)
Non-binary (%)	-	1 (3.6)

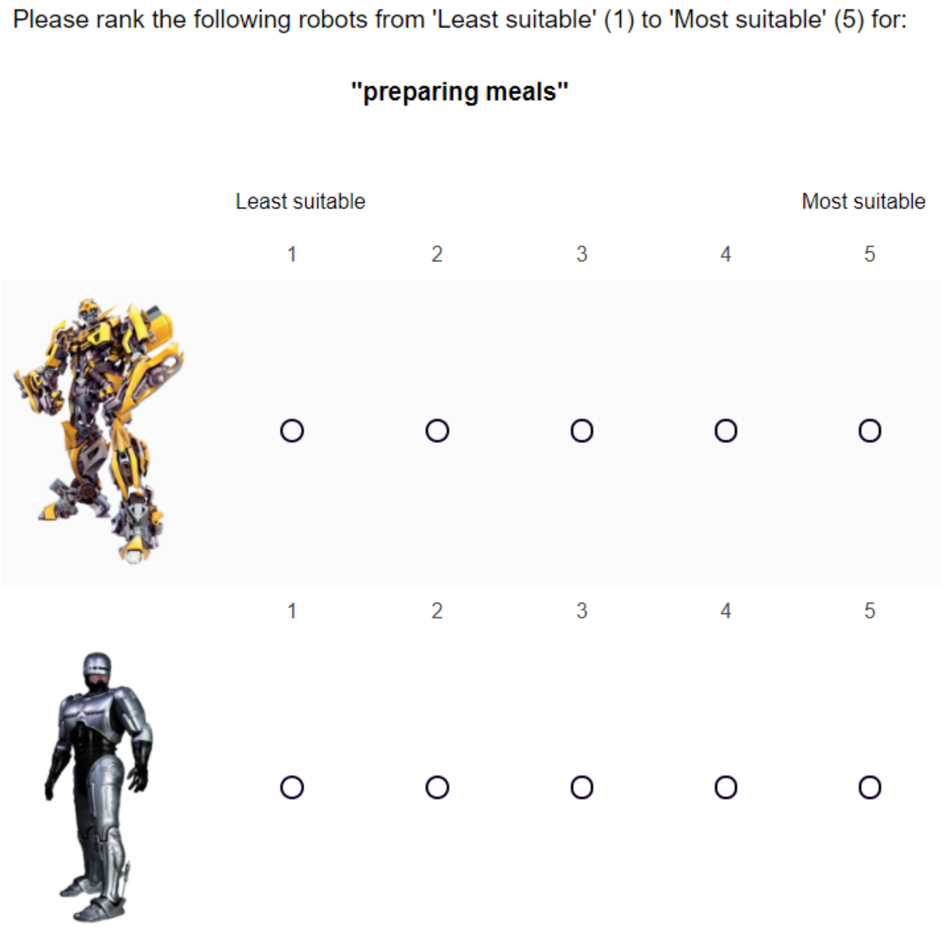
5.4.1.2 Materials and Procedure

Participants were randomly assigned to one of two conditions automatically via Qualtrics' randomisation function. After participants' consent was obtained, they were asked to undertake a task designed to prime either fictional or non-fictional representations of robots (depending on condition) before completing measures of their implicit and explicit attitudes toward robots.

[16] A priori power analysis was not conducted for this study due to an oversight by the author.

Figure 5.3

Example of a Part of the Priming Task



Priming task. In an attempt to prime participants with different representations of robots, a ranking task was constructed for the purposes of this study (see Figure 5.3). Participants were shown five fictional or non-fictional robots (selected in the Pilot Study; see Section 5.3.2) at the same time and asked to rank the robots from 1 (*least suitable*) to 5 (*most suitable*) for performing five different tasks (preparing meals, cleaning a house, taking care of an elderly person, performing surgery, and teaching children to read). The robots were presented multiple times and in random order for each of the tasks.

Implicit Association Test (IAT). Participants were then asked to complete MacDorman et al.'s (2009) robot-human IAT was adapted following the guidance provided by Greenwald et al. (2003) to assess participants' implicit attitudes toward robots, relative to humans. For a full description of the task, see Section 5.2.1.2.

Negative Attitudes toward Robots Scale (NARS). Following the IAT, participants were asked to complete a modified version of the Negative Attitudes toward Robots Scale (NARS) to measure their explicit attitudes toward robots (Nomura et al., 2004). For a full description of the NARS, see Section 5.2.1.2.

Manipulation checks. Since Thellman and Ziemke's study suggests that responses to the NARS can be influenced by presenting participants with specific visual examples of robots, an open-ended question was presented to participant immediately afterward the NARS in order to explore whether participants had any particular robots or experiences in mind whilst completing the NARS ("Did you have any particular robot(s) or experience(s) in mind when you were answering the questions on the previous page?"). This question was intended as a way to gain some additional insight into what representations of robots people may have in mind when completing measures of their attitudes toward robots.

Participants were also asked an open-ended questions used to assess whether participants were likely to be naïve to the experimental manipulation ("Did you notice any similarities between the robots you had to rank? If yes, in what way were they similar?"). Participants were then asked to rate each of the robots they saw (either fictional or non-fictional) on the seven dimensions used in Pilot Study 4 (see Section 5.3.1); *pleasant – unpleasant, hostile – friendly, safe – threatening, unfamiliar – familiar, real – fictional, evil – good, humanlike – mechanical*. This was done in order to (a) check that participants correctly identified the robots as fictional and non-fictional, and (b) assess whether there were any other differences between the images of the fictional and non-fictional robots. Finally, participants were asked to provide their age and gender, entered into the prize draw (or awarded credits if students), and debriefed (see Appendix D for debrief procedure).

5.4.2 Results

Raw data was extracted from Qualtrics and copied into an Excel workbook. All analyses were conducted using SPSS. Three participants identified the fictional status of the robot images in the fictional condition as a similarity between the robots. However, these participants were not excluded as they were: (a) not identified as statistical outliers and (b) their exclusion did not result in any statistically meaningful changes to the results from the analyses described below. Two separate analyses are conducted and reported below. A one-way MANOVA without the inclusion of any covariates and a one-way MANCOVA with a single covariate, goodness, to account for difference between fictional and non-fictional robots that were not part of the manipulation. Treatment of the data and assumption checks relevant to the analyses carried out in this section can be found in Appendix Q.

5.4.2.1 Analysis without covariates

Effects of priming on participants' explicit and implicit attitudes. A one-way MANOVA was conducted to determine whether presenting (i.e., priming) participants with images of fictional or non-fictional robots had an effect on their subsequently measured explicit and implicit attitudes toward robots (see Table 5.23). No effect of the experimental manipulation on participants' NARS and IAT scores was found, Wilks' Λ = 0.88, $F(4, 51) = 1.71$, $p = .162$, $\eta^2_\Lambda = 0.12$.[17]

Table 5.23

Descriptive Statistics for Each Dependent Variable and for Each Condition

Measure	Condition	*M*	*SD*
IAT	Fictional	0.64	0.36
	Non-fictional	0.50	0.37
NARS-S1	Fictional	13.00	3.02
	Non-fictional	12.43	4.09
NARS-S2	Fictional	16.36	3.66
	Non-fictional	15.82	4.25
NARS-S3	Fictional	14.43	4.11
	Non-fictional	15.18	5.03

[17] Note that effect size for Wilks' Λ has been estimated following the recommendations of Steyn Jr and Ellis (2009) and Olejnik and Algina (2000), using the formula $\eta^2_\Lambda = 1 - \Lambda$ where .02 is a small effect, .13 is a medium effect, and .26 is a large effect.

5.4.2.2 Ratings of the robot images

Participants ratings of the robots (fictional or non-fictional) that they saw during the study were explored using a one-way MANOVA to determine whether there were significant differences in the mean ratings of the images between the two conditions that would need to be controlled for in the main analysis. The mean rating that each participant gave for all seven dimensions was calculated by averaging the ratings of the five robots (fictional or non-fictional) that participants viewed during the priming task. Assumptions were tested using the same procedure as the one described in Section 5.2.2. A significant difference was found between the two conditions as indicated by Pillai's trace, $V = 0.70$, $F(7, 48) = 16.22$, $p < .001$.[18] This result was followed up using univariate one-way ANOVAs in order to identify on which dimensions participants' ratings differed in the two condition (see Table 5.24). A Bonferroni correction was applied for $n = 7$ comparisons resulting in an adjusted critical α level of $p < .007$. Although the Bonferroni adjustment is not divergent from common correction practices for multiple comparisons, its use can be questionable especially when the number of comparisons is relatively large, thus increasing the risk of Type II errors (Cabin & Mitchell, 2000). Since no suitable alternative to this particular adjustment was found in this case, 95% CI are also reported in Table 4 to aid in the interpretation of the results.

[18] Note that Pillai's trace was used as Box's Test of Equality was significant and V is considered more robust to this type of violation (Pillai, 1955).

Table 5.24

Descriptive Statistics and t-tests for Each of the Seven Dimensions

Dimension	Fictional robots		Non-fictional robots			
	M (SD)	95% CI [LL, UL]	*M (SD)*	95% CI [LL, UL]	*F*(1, 54)	*p*
P – U	3.11 (0.81)	[2.77, 3.46]	3.62 (1.00)	[3.28, 3.97]	4.36	.042
H – F	4.66 (0.72)	[4.32, 5.01]	4.49 (1.05)	[4.15, 4.83]	0.51	.480
S – T	3.29 (0.83)	[2.94, 3.63]	3.55 (1.00)	[3.20, 3.90]	1.16	.287
U – F	4.97 (1.33)	[4.54, 5.41]	3.85 (0.94)	[3.42, 4.29]	13.36	.001*
R – F	5.89 (1.00)	[5.49, 6.30]	3.33 (1.12)	[2.93, 3.73]	81.26	.000*
E – G	5.20 (0.69)	[4.91, 5.49]	4.34 (0.84)	[4.04, 4.63]	17.60	.000*
H – M	4.57 (0.90)	[4.20, 4.95]	4.77 (1.08)	[4.40, 5.15]	0.57	.454

Note. Each dimension is on a scale from 1 to 7 with the most extreme value for the first word always being 1. P-U denotes the pleasant-unpleasant dimension; H-F denotes the hostile friendly dimension; S-T denotes the safe-threatening dimension; U-F denotes the unfamiliar-familiar dimension; E-G denotes the evil-good dimension; H-M denotes the humanlike-mechanical dimension.

* Significant at $p \leq .007$.

A significant difference between the ratings that participants gave the images of the fictional and non-fictional robots on the unfamiliar – familiar, real – fictional, and evil – good dimensions. As expected, the fictional robots were rated as more fictional than the non-fictional robots which confirms that participants were able to identify the types of robots (fictional or non-fictional) that they were exposed to. This particular dimension served as a way to check whether the fictional and non-fictional robots were recognised as such by participants. It is therefore not something we would wish to control for as it is part of the experimental manipulation.

Unfortunately, but not entirely unsurprising given the difficulty with matching the robot images during the pilot study, participants rated the fictional robots as significantly more familiar and good than the non-fictional robots. Since it was not possible to control for these differences via the selection of matching images, it was necessary to consider including these dimensions as covariates in the model in order to control for them.

5.4.2.3 Selection of covariates

As discussed in the section above, participants' ratings of the robots on the evil-good and unfamiliar-familiar dimensions were significantly different between the two conditions which was not an intended part of the experimental manipulation. Three criteria were used to assess whether these variables would be suitable as covariates. Covariates must: (a) be statistically different between conditions but not be an intended part of the experimental manipulation; (b) be associated (i.e., correlated) with the dependant variables, but not directly manipulated and likely to introduce additional variance; and (c) not substantially overlap with other covariates (i.e., there must not be a significant statistical relationship between them).

Both the evil-good and unfamiliar-familiar ratings satisfy the first (i.e., they are both not an intended part of the manipulation but significantly differ between the conditions) and the second criterion (i.e., they are both correlated with the dependent variables as presented in Table 5.25 and Table 5.26). However, these two dimensions do overlap. There is a significant relationship between participants' evil-good and unfamiliar-familiar ratings such that more unfamiliar robots are also, in general, rated as more evil (see Table 5.25 and Table 5.26). As such, a decision was made to include only one of these dimensions as a covariate. Since the evil-good dimension was, in general, more strongly associated with the dependant variables, it was decided that it would be entered into the analysis as a covariate instead of the unfamiliar-familiar dimension.

Table 5.25

Correlations (Pearson's r) Between the Dependant Variables and Possible Covariates for the Fictional Condition

Variables	IAT	S1	S2	S3	Age	U-F	E-G
IAT	-	-	-	-	-	-	-
NARS-S1	-.01		-	-	-	-	-
NARS-S2	.31	.55**	-	-	-	-	-
NARS-S3	.31	.66**	.66**	-	-	-	-
Age	-.04	-.33	-.35	-.29	-	-	-
Unfamiliar – Familiar	-.28	-.18	-.19	-.38*	.15	-	-
Evil – Good	-.63**	-.09	-.34	-.34	.08	.56**	-
Real – Fictional	-.26	-.46*	-.27	-.42*	.27	.68**	.41*

Note. NARS-S1: Interaction with robots; NARS-S2: Social influence of robots; NARS-S3: Emotion when interacting with robots, IAT: d-scores.

*$p \leq .05$;

** $p \leq .01$.

Table 5.26

Correlations (Pearson's r) Between the Dependant Variables and Possible Covariates for the Non-Fictional Condition

Variables	IAT	S1	S2	S3	Age	U-F	E-G
IAT	-	-	-	-	-	-	-
NARS-S1	.34	-	-	-	-	-	-
NARS-S2	.39*	.70**	-	-	-	-	-
NARS-S3	.41*	.80**	.74**	-	-	-	-
Age	-.07	-.37	-.41*	-.32	-	-	-
Unfamiliar – Familiar	-.08	-.25	-.27	-.28	.36	-	-
Evil – Good	-.18	-.57**	-.50**	-.55**	.43*	.52**	-
Real – Fictional	-.03	.18	-.05	.07	-.36	-.18	.02

Note. NARS-S1: Interaction with robots; NARS-S2: Social influence of robots; NARS-S3: Emotion when interacting with robots, IAT: d-scores.

*$p \leq .05$;

** $p \leq .01$.

5.4.2.4 Analysis controlling for differences between fictional and non-fictional robots

Evaluation of assumptions. The majority of evaluation has been completed in the section of the previous analysis and additional assumptions for a one-way MANCOVA were evaluated and met.

Effects of priming on participants' explicit and implicit attitudes. A one-way MANCOVA was conducted to determine whether presenting (i.e., priming) participants with images of fictional or non-fictional robots had an effect on their explicit and implicit attitudes toward robots (see Table 5.27 for adjusted means) when controlling for the evil-good rating given to the robots. There was a significant difference between participants' explicit and implicit attitudes between the two conditions when controlling for the robots' goodness, Wilks' Λ = .80, $F(4, 50) = 3.22$, $p = .020$.

Table 5.27

Adjusted Means and Standard Errors for Each of the Dependent Variables

Measure	**Condition**	M_{adj}	***SE***
IAT	Fictional	0.72	0.07
	Non-fictional	0.42	0.07
NARS-S1	Fictional	13.78	0.68
	Non-fictional	11.64	0.68
NARS-S2	Fictional	17.33	0.74
	Non-fictional	14.85	0.74
NARS-S3	Fictional	15.63	0.84
	Non-fictional	13.98	0.84

Bonferroni corrected (adjusted critical α level of $p < .013$) one-way ANCOVAs were used to follow-up the significant multivariate effect. There was a significant difference in the adjusted means of the IAT d-scores between the fictional (IAT = 0.72) and non-fictional condition (IAT = 0.42), $F(1, 53) = 7.84$, $p = .007$, partial $\eta^2 = .13$. It was found that participants exposed to non-fictional representations of robots had less negative implicit attitudes toward robots than participants exposed to fictional representations of robots. There was no significant difference in the adjusted means of the Interaction (F(1, 53) = 4.34, p = .042, partial η^2 = .08), Social influence (F(1, 53) = 4.99, p = .030, partial η^2 = .09), and Emotion (F(1, 53) = 1.70, p = .199, partial η^2 = .03) subscales of the NARS between the fictional and non-fictional conditions.

5.4.3 Discussion

Based on the findings from Thellman and Ziemke (2017), it was expected that there would be a significant difference in participants' implicit and explicit attitudes depending on whether they were shown images of fictional or non-fictional robots. The hypothesis was partially confirmed as a significant difference in people's implicit (but not explicit) attitudes was found between the two conditions. While participants in both conditions associated positive words more strongly with humans than with robots, participants who were primed with non-fictional representation of robots were less negatively biased toward them (i.e., they had less negative implicit attitudes toward robots). Although the direction of the difference is consistent with that of Study 5 (where robots presented as non-fictional were more positively evaluated), it is unclear what the implications of Study 6 are or why only implicit attitudes were affected. One possible explanation could be that images of the non-fictional robots were more similar to the stimuli (i.e., silhouettes) used in the IAT to represent the category of robots. As such, showing participants the images of the non-fictional robots primed them to respond more quickly during the IAT than the participant who were shown images of the fictional robots. This explanation is supported by the relatively small difference in IAT scores between the two conditions as well as the fact that, on average, participants in both conditions associated positive words more strongly with humans than with robots. Unfortunately, it is not possible to say whether this explanation hold true without knowing how similar the images of real robots were to those used in the IAT.

Although no significant difference was found between conditions for participants' explicit negative attitudes, there were some general findings that are potentially of some interest. Overall, participants' attitudes were relatively positive in relation to the first subscale of the NARS, interaction with robots. This may be indicative of participants' general openness to interact with robots or a reflection of their past human-robot interaction experiences (if any). Since no attempt was made to measure participants' direct experience with real robots, it is unclear to what extent this may have affected participants' responses to this particular measure of attitudes. It should be noted that two participants stated that while answering the NARS they kept in mind specific real-life interaction experiences with robots, as well as four participants who specifically mentioned thinking about non-fictional robots. With reference to the findings from Study 4, it could be that recalling experiences with or knowledge of real robots is particularly impactful on people's attitudes and may negate experiences with fictional robots. However, it is not clear how such experiences might have interacted with the IAT and the images with which participants were primed.

In regards to the lack of significant differences of participants' explicit negative attitudes toward robots between the conditions, there are a number of possible explanations. It could be that priming participants with different representations of robots simply has no effect because participants' attitudes are not influenced by whether robots are fictional or not, and that fictional representations of robots play no role in shaping people's attitudes toward robots. This is, of course, impossible to confirm in the current study and maybe unlikely given that we found a significant difference of people's implicit attitudes between the two conditions. An explanation that is somewhat more likely is that the priming task was either not sufficiently strong to affect participants' explicit attitudes toward robots or was in some way limited given the choice of images (as discussed above). As discussed previously, it is likely that at least some participants had particular experiences and/or robots in mind when filling out the NARS. In fact, three participants went into some detail explaining their concerns in regards to robotics and artificial intelligence. Therefore, it is somewhat questionable whether presenting participants with images of robots in a limited context (i.e., participants preference for the design of the robots for particular tasks) would be enough to influence participants who already have strong views about robots and their use. However, it should also be noted that less than a quarter of participants stated that they had any experiences or robots in mind when completing the NARS and that, given the limited data, it is difficult to establish whether there is any relationship between those statements and participants' explicit negative attitudes. Of course, it is also possible that some participants either chose not to report any experiences or robots they had in mind while completing the NARS. Overall, Study 6 was not particularly helpful in understanding the impact of fictional depictions of robots on people's attitudes although it does demonstrate that fictional and non-fictional robots are perceived differently on a number of dimensions. Avenues for future research are discussed in Chapter 6.

Chapter 6: Discussion

Chapter Summary

This chapter will provide an overview of the empirical work presented in this thesis with respect to the three main questions this thesis set out to answer: (a) What representations typically come to mind when people think about robots?; (b) Are these representations linked to people's attitudes toward robots?; and (c) Do fictional portrayals of robots influence how people represent robots and people's beliefs about, and attitudes toward, robots? Some of the more notable limitations of the work and future research directions will also be discussed.

6.1 What representations typically come to mind when people think about robots?

The systematic review presented in Chapter 2 supported the previously observed variability in people's attitudes toward robots (Takayama et al., 2008; Enz et al., 2011; European Commission, 2012). Much like previous work, the findings suggest that such variability can be at least partially explained by a number of factors such as the domain of application of a robot, the type of human-robot interaction, and people's individual characteristics (e.g., nationality). However, these factors alone neither account for all of the observed variability in people's attitudes nor do they explain why and how these factors shape said attitudes. Furthermore, the general decline in favourable attitudes toward robots among EU citizens that has been observed in recent years (Gnambs & Appel, 2019) also can not entirely be explained by looking at individual factors. This thesis therefore took a different approach by considering whether the way people internally represent robots may account for the observed variability in attitudes and how fictional (and non-fictional) media may influence said representations.

The studies in Chapter 3 drew on two theoretical frameworks by Abric (1993) and Lord and Lepper (1999) in order to investigate how people think about robots. Study 3 presented evidence on the social representation of robots – that is, how a given group (in this case, broadly defined as the general public in the UK) defines and thinks about another group or category (in this case, robots). A social representation should be viewed as the 'average' or 'typical' representation of robots rather than an individualistic representation (Abric, 1993). As such, Study 3 focused on Abric's conceptualisation of social representations and aimed to answer the first question posed in this thesis – namely, what representations typically come to mind when people think about robots? Study 3 expanded upon Study 2 as well as the previous, but ultimately limited, work of Piçarra et al. (2016) in order to elaborate upon the social representation of robots. This was done by using a more sophisticated analytic approach in the form of a semantic network that utilised an algorithm for detecting communities (i.e., structures) in such networks. The core of the representation was reflected by 10 associations (called *global hubs*) that had a large number of connections to other elements in the network. Overall, Study 3 suggested that, at the core, robots are viewed as essentially artificial creations that, although apparently intelligent and useful, are emotionless and somewhat removed from the social domain.

Initially, this finding might suggest that robots are not a novel and distinct social group that is forming as a result of the advent of social robots as has been suggested in recent years (Vanman & Kappas, 2019). Alternatively, it could be that more advanced social robots have not been present nearly as long as manufacturing robots or more crude, non-humanoid robots (Bartneck, 2004; Behnke, 2008). Arguably this lack of presence of socially competent robots is true of both real-life and fiction; and with an average age of 30 years for the sample in Study 3, it is not surprising that the core of the representation seems to be closer to what may be considered a more traditional portrayal of robots (Kriz et al., 2010). Additionally, participants in Studies 2 and 3 were asked to make associations with the word "robots" and not with "social robots", which suggests that it may be valuable to explore the differences between the social representation of different types of robots in future research. For example, via the construction and comparison of semantic network for both concepts in a manner comparable to the one used by Keczer et al. (2016).

Although the core of the social representation is important as it is the basis upon which the representation is built, it may not necessarily contribute to the variability in people's attitudes toward robots (see Chapter 3, Section 3.1). Instead, it is the peripheral elements surrounding the core that are less homogeneous and likely reflect the variety of people's experiences and attitudes. In Study 3, every association that was not a global hub in theory represented the periphery of the representation. The method used by Keczer et al. allowed for the detection of communities (or related elements) in the semantic network that showed how the periphery of the representation was structured around the core in a number of modules. The emergence of a modular structure for the semantic network indicated that the way that robots are represented across a predominantly UK based sample can be divided into five distinct communities that reflect the different ways in which robots are conceptualised: (i) the core characteristics of robots based on fictional and historical shared experience; (ii) the potential usefulness of robots; (iii) robots as emotionless but efficient; (iv) robots as artificial creations; and lastly (v) robots in their role as machines. The meaning of these modules was inferred from the associations that were related to the core, the average valence of each module, and a broad look at the peripheral associations. As such, the meaning of the modules was largely open to interpretation, especially given the absence of research on the topic. Study 4 did, however, support the interpretations of the five modules, as the themes that emerged from the qualitative analysis in Study 4 were largely synonymous with the conclusions drawn from Study 3. Overall, Study 3 extends current understanding of how people represent robots and demonstrates the variability in the way robots are conceptualised. Additionally, Study 3 offered a novel approach to investigating the representation of robots that can be replicated in the future with similar or different samples thus allowing for the comparison of representations across time and cultures.

6.2 Are individuals' representations of robots linked to their attitudes?

In addition to identifying how people think about robots, Study 4 also considered the relationship between the way people represent the concept of robots and their explicit attitudes. As such, Study 4 primarily informed the second research question in this thesis which was inspired by Lord and Lepper's (1999) Attitude Representation Theory.

There was no evidence that endorsement of the association between robots and "artificial intelligence and computing" contributed to the variability in people's attitudes as all participants endorsed the association. This finding is consistent with the conclusions drawn from Study 3 and the characteristics of the core as described by Abric. In other words, the core is a relatively stable construct that is unlikely to be influenced by novel experiences with robots and thus contribute to changes in attitudes. Similarly, there was no strong evidence that endorsing associations between robots and "useful and helpful" and "artificial" contributed to the variability in people's attitudes. This finding suggested that not only do most people recognise robots as currently and/or potentially useful but that doing so does not necessarily predict whether that individual will have strongly positive attitudes toward robots. It is however possible that someone who feels strongly that robots are not (or will not) be useful may hold more negative attitudes, especially toward the use of robots. Unfortunately, as only two participants did not endorse the association with "useful and helpful", it was not possible to ascertain how likely this explanation is. There is some evidence that the perceived usefulness of specific robotic systems is linked to people's intention to use said robots (De Graaf & Allouch, 2013; Rantanen et al., 2018) which is a question that future research could investigate further with reference to the findings from Study 4. For example, conducting additional semi-structures interviews focused specifically on understanding how representing robots as useful (or not) impacts people's general attitudes especially when considering the different ways in which people see robots as useful (e.g., useful for manual labour and nothing else).

In regards to the association with "artificial", the findings suggested that viewing robots as artificial or not was not only unrelated to people's attitudes but also rather heterogeneous in terms of participants' understanding of artificiality. There was, however, evidence that the modules represented by the association with "emotionless and cold" and "machines" were related to participants' attitudes in a predictable manner. Participants who did not endorse an association between robots and machines had slightly more positive attitudes toward interacting with robots and the potential social influence of robots than participants who did not associate robots as machines. The qualitative analysis revealed that robots are generally perceived as distinct from machines; although some participants also indicated that some robots are perceived as less or more machine-like depending on their function and appearance. This finding suggests that individuals who perceive robots as different from machines, especially if they conveyed positive sentiment regarding robots as a separate entity, also had more positive attitudes toward interacting with robots in the future. Much like artificiality, participants' discourse surrounding machines was somewhat heterogeneous. For example, some participants appeared to have negative feelings toward machines while others felt more comfortable with them. With reference to Lord and Lepper's theory, it may be expected that individuals who hold representations of robots that are centred around their role as machines may be less favourable toward social robots as that would conflict (i.e., not *match*) with their conceptualisation of robots.

The clearest finding was that participants who endorsed an association between robots and "emotionless and cold" had more negative attitudes toward robots. This difference was most notable for the third NARS subscale which measured participants' affective attitudes toward prospective interaction with robots. This finding suggests that viewing robots as emotionless entities likely predisposes individuals to be sceptical about the extent to which interacting with robots will be a positive experience. However, endorsement alone is not enough to explain some of the subtler variation among participants' attitudes. For example, the qualitative analysis suggested that, while some participants envisioned how robots with emotion could be useful and ultimately desirable, others rejected the idea for various reasons. In other words, participants who were more open to encountering emotion in robots (be it genuine or programmed) also had overall more positive attitudes toward robots. This was true regardless of whether participants endorsed the idea that robots are typically emotionless which may have been in part due to the fact that some participants' personal experiences consisted of positive interactions with robots. Multiple studies have demonstrated that direct interaction with social robots generally leads to more positive attitudes toward robots (Bartneck et al., 2007; Wullenkord et al., 2016). It could be that where social robots are concerned, experiencing an interaction that portrays robots as less emotionless (e.g., a robot programmed to express various emotions) may lead to more positive attitudes toward emotion in robots. Interestingly, the systematic review in Chapter 2 found that people tended to have 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. It is unclear what the implications of this finding are and may warrant further investigation. For example, an experimental study comparing direct (face-to-face) and indirect (watching a video) interaction with a social robot expressing varying degrees of emotion (e.g., simulated facial expressions).

Overall, the studies presented in Chapter 3 identify how people represent robots and suggest that the variability in attitudes toward social robots may be partially explained by some of the ways in which robots are represented. This conclusion supports Attitude Representation Theory (Lord & Lepper, 1999) as it appears likely that people's attitudes at a given time reflect what individuals have in mind at the moment that they are asked about their attitudes. There are practical implications for survey research following on from the work presented in Chapter 3. First, it is clear that caution must be used when asking questions about robots as a broad category - especially where attitudes are concerned – as people think of robots in diverse and often contradictory ways. Where possible, visual or textual examples of robots should be provided in order to make sure that participants are guided to make evaluations congruent with the researchers' own understanding of robots. Additionally, it may be helpful for researchers to interrogate their own internal representation of robots (e.g., via the five modules described above) and think about how and why their understanding of robots may differ from that of the general population.

6.3 The role of fiction in the social and individual representation of robots

All three studies presented in Chapter 3 support evidence (DiSalvo et al., 2002; Riek et al., 2011; Mubin et al., 2015) that fictional robots are at the forefront of people's minds when they engage with robots directly or indirectly. There are two possible explanations for this finding: (i) people are more familiar with fictional robots, and / or; (ii) fictional portrayals of robots are more memorable in some way and thus more likely to be at the forefront of people's minds. Participants in both Studies 4 and 5 reported more experiences relating to fictional robots than real robots which gives some weight to the first explanation and implies that people's representations of robots may be primarily based upon fiction. However, not everyone engages with fiction in the same way and to the same extent (Green, 2004; Busselle & Bilandzic, 2008) which begs the question of whether fictional representations of robots are so pervasive in popular culture that their salience is a result of their presence rather than signifying deeper engagement with fictional portrayals. The findings of Study 3 also put into question whether the salience of fictional robots reflects their impact as there was no relationship between the salience of fictional robots and people's explicit and implicit attitudes. This finding was somewhat surprising as it was expected that such a relationship would exist as Riek et al. (2011) found a positive correlation between the number of fictional films that people had watched and their attitudes. However, being able to name fictional robots does not necessarily equate to engagement with fictional media as some robots (e.g., R2D2 from the Star Wars franchise) are known even to people who are not familiar with the depictions.

It could be that fictional portrayals are interpreted in different ways by different people. This idea is supported by the findings of Study 4 as it was observed that, while for some participants' negative portrayals of robots contributed to their concerns surrounding robotics, other participants acknowledged these portrayals but were sceptical that they would occur in real-life and as such did not express explicitly negative sentiment. In general, the work presented in this thesis suggests that participants with more direct experiences with robots also had more positive attitudes which may have enabled them to disconnect from (negative) portrayals of robots in fiction. This finding is supported by research on the way people process fiction based on their own experiences. For example, Prentice et al. (1997) found that students who read fictional stories which contained dubious assertions about the real world were more likely to believe the assertions if the fictional story was set in a realistic but unfamiliar setting as opposed to a familiar one. In the case of robots, the setting of a fictional narrative may be less important but familiarity with real robots may mean that people identify and question depictions of robots which are dissimilar to their lived reality.

The qualitative analysis for Study 4 suggested that for all associations, fictional portrayals informed participants' concerns and the consequences that they imagined might result from the advent of robots. This role of fiction was most notable in regards to the association between robots and "artificial intelligence" and "emotionless" as participants generally recognised the link between these associations and fictional depictions of robots (e.g., robots depicted as lacking emotion). Given that endorsing or not endorsing the association between robots and "emotionless" was linked to participants' attitudes, the role of fiction in strengthening that association may be of particular significance. This is especially true if we consider Kriz et al.'s (2010) finding that the majority of robots in recent films were portrayed as having good (or at least superior-to-human) cognitive abilities (e.g., problem solving) but generally lacked human-like social behaviours that may contribute to robots being perceived as emotionless and cold. However, the relationship between fiction and the endorsement of robots as emotionless was not as straightforward. Participants who reported more direct experiences with robots, and ultimately viewed robots as friendly, were less likely to speak of robots as emotionless in an overtly negative way even if they endorsed the association and regardless of their experience with fictional robots. As such, it is likely that direct experience with real robots can counteract typical depiction of robots in fiction (such as the ones described by Kriz et al.).

In all three studies reported in Chapter 3, some participants mentioned both fictional and non-fictional robots when asked to retrieve such examples which questions the second explanation for the salience of fictional representations; namely, that fictional robots are more memorable. Although, based on the findings from Study 4, it is likely that novel experiences with fictional or non-fictional robots play a significant role in remembering specific robots and perhaps in the formation of people's attitudes. Some notable examples from the participants in Study 4 were first encounters with real robots (generally positive and in one case, disappointing) and childhood encounters with fictional robots that, according to participants, influenced how they thought and felt about robots. In light of the findings by Prentice et al. (1997) presented earlier on in this section, future research should investigate how people's personal experiences with robots lead them to interpret fictional depictions of robots and which aspects of those depictions they are likely to incorporate into their personal representation of robots. For example, future research may ask participants to interact with a social robot and then be divided into groups to watch fictional robots taking part in comparable positive or negative interactions. Participants could then be asked to report their attitudes toward robots in general and asked to evaluate the fictional interactions they watched. Their responses could be compared to the attitudes and evaluations of a control group that did not take part in the direct interaction with a robot.

Overall, the evidence presented in Chapter 3 supports the idea that fictional portrayals influence people's understanding and representation of robots but that the degree to which they do so may be depend on other factors (e.g., direct experience with real robots, specific aspect of the representation).

6.4 The effect of fictional and non-fictional portrayals of robots on people's attitudes

Chapter 5 presented two experimental studies that investigated whether making fictional versus non-fictional portrayals of robots salient influenced people's attitudes. Study 5 found that participants who were presented with information describing a video and the robots contained within as non-fictional had more positive attitudes toward using the robots depicted in the video, as well as more positive attitudes toward interacting with robots in general, and more positive attitudes toward the affective outcome of interaction with robots in general. The findings of this study are somewhat in line with findings by Mubin et al. (2015) who observed that YouTube videos of two non-fictional robots (Nao and Shakey) generated more engagement and positive interest (i.e., positive comments) compared to videos of fictional robots (HAL900 and Astro Boy). Fictional characters (including robots) may be more prone to negative comments due to their affiliation with particular sci-fi movies and their role within them. Alternatively, it could be that non-fictional robots are less familiar to viewers and, as such, are more novel, thus leading to more engagement with videos depicting non-fictional robots. Although pre-existing associations may explain the findings of Mubin et al.'s study where the fictionality of the robots was not manipulated, they do not necessarily explain why identical robots that were unfamiliar to participants (as used in Study 5) would be subject to such biases.

One possible explanation could be that labelling the robots as non-fictional and convincing participants that these robots were currently in use, allowed participants to more easily imagine using said robots, which resulted in more positive attitudes toward their use. If this is indeed the case, it may have wider implications in regards to non-fictional material accessible by the general public, especially where the abilities and availability of more advanced robots could be exaggerated or misleading. For example, robot demonstrations featured on talk shows or public engagement events. Such exaggerations could be especially impactful as there was no evidence that participants who believed the manipulation (regardless of the condition to which they were assigned) had less experience with fictional and non-fictional robots than participants who did not believe the manipulation. It is likely then, that if the setting and robots are plausible and realistic, members of the general public may not be able to assess whether the robots are realistic, even under heightened scrutiny of non-fictional material (Green et al., 2006). A practical implication is that researchers may need to be mindful of the ways they demonstrate the abilities of robots and seek to be as transparent as possible about how robots work and what they can and cannot do.

It is not clear why being led to believe that the video depicts non-fictional rather than fictional robots would result in more positive attitudes toward robots in general. Although participants were asked if they had anything in mind while answering the

NARS, the vast majority reported that they did not think or imagine anything while answering the questions. That does not however mean participants did not rely on their individual representation of robots which, as Chapter 3 indicates, likely influences people's attitudes toward robots. It could be that the video and robots are more prototypical representations of non-fictional rather than fictional robots. In other words, participants in the fictional condition may have had less positive attitudes toward robots because what they observed was not a typical robot that may appear in fiction. Indeed, as described in the results, participants, on average, reported more experiences with fictional robots than with non-fictional robots, meaning that they were less likely to have in mind a prototypical model of a real robot to which they could compare the robots in the video. Since the typicality of the robots in the video was not evaluated, this possibility cannot be discounted.

6.5 Limitations and future directions

Existing research on attitudes toward robots – including the work presented in this thesis - face the same problems in terms of the measurement of attitudes as other fields. Namely, a notable reliance on self-report methods, especially the NARS (see Chapter 2, Sections 2.3 and 2.4.5), as well as general concerns about the reliability and validity of some self-report measures. While using a well-known and validated measure such as the NARS allows for comparison between studies and is the primary measurement of attitudes used throughout this thesis, there are a number of issues in the way the NARS is typically used. When talking about the variability in people's attitudes toward robots, it may be more apt to say that this thesis examined the variability of participants' responses to the three NARS subscales. Although often utilised as a measure of people's attitudes toward robots in general, the NARS (Nomura et al., 2004) arguably contains a larger number of questions measuring affective rather than cognitive attitudes which is generally not acknowledged in studies, especially when the NARS is not divided into its subsequent subscales. As noted in the systematic review presented in Chapter 2, it is possible for people to think differently about robots than how they feel about them (see Section 2.4.1). This discrepancy is highlighted in Study 4 where some participants felt strongly about specific topics (e.g., sentience in robots is scary) but simultaneously acknowledged that it did not necessarily change how they thought about robots (e.g., sentience not possible, so not a real concern). Moreover, the NARS contains questions relating to robot sentience and / or emotions in all of its subscales which were more sensitive to the variability of participants' attitudes with reference to the themes that emerged from Study 4. Regardless of whether the NARS or other self-report measures are used, future work may benefit from a closer examination of measures that may be susceptible to capturing a discrepancy between people's feelings and thoughts about robots. With respect to this thesis, Study 4 does suggest that fiction may be more impactful in terms of people's feelings toward robots which may not necessarily be congruent with their knowledge and experience of real robots or change what people think about the use of robots. However, this was not considered outside of Study 4 and presents an interesting avenue for investigation in the future.

Although the inclusion of the human-robot IAT developed by MacDorman et al. (2009; based on Greenwald et al., 1998) was intended to provide an alternative measure of attitudes for some of the studies in this thesis, it too has problems. A number of critiques have been made regarding the IAT's reliability and validity since its conception (Fiedler et al., 2006). The test-retest reliability of the IAT ($r \approx 0.50$; Norsek et al., 2007) has been particularly criticised for not reaching the ideal standard of $r > 0.70$. Although this concern in not exclusive to the IAT, it does have implications for measuring changes over time and changes as a result of experimental manipulation (which was the case for the studies presented in Chapter 5). One particular criticism of the IAT is of great relevance to the work presented in this thesis - namely how useful it is as a measure of differences between individuals. According to a relatively recent review of the criticisms surrounding the IAT, Meissner et al. (2019) point out that the IAT is susceptible to the influence of a number of individual factors (e.g., familiarity with the IAT) which means that the IAT is not a pure measure of the differences in people's implicit attitudes. Individual factors were not considered in this thesis and add to the limited usefulness of the IAT in relation to the studies presented in Chapter 5. Furthermore, this critique extends to a broader question about the usefulness of the IAT as a predictor of behaviour as multiple studies have demonstrated that the IAT does not reliably predict behaviour (Meissner et al., 2019). This has been especially important to the discourse surrounding the use of the IAT to predict discriminatory behaviour toward outgroup members (Oswal et al., 2015), which may bear relevance to the study of human-robot interaction if robots are perceived as a social outgroup.

Additionally, the stimuli used for the human-robot IAT do not appear to have been piloted and it is unclear how representative the robot silhouettes are of robots as a broad category. Given that Chapter 3 demonstrated the variety of ways in which people represent robots, a more rigorous selection and evaluation of the stimuli for the IAT may be of benefit for future research. For example, the stimuli can be evaluated by a diverse sample made-up of the general public based on a number of dimensions such as how recognisable as robots the images are, how human-like and machine-like the images are, and so on. Furthermore, it could be said that the human-robot IAT more accurately captures bias for humans over robots rather than attitudes toward robots in general. It is therefore important to consider other ways of measuring attitudes that do not rely wholly on self-report and circumvent the major shortfall of the IAT. For example, the single category IAT developed by Karpinski and Steinman (2006) uses a single attitude object (e.g., robot) rather than two objects (e.g., robot and human) and as such may be a suitable alternative to MacDorman et al.'s human-robot IAT. Although, given the above stated limitations of the IAT as a measure of attitudes, there is arguably a need to develop alternative measures of attitudes toward robots.

The work presented in this thesis is not immune from the criticisms of the wider body of work discussed in Chapter 2 (see Section 2.4.5). More specifically, the size and composition of participant samples. Although the studies presented in the preceding chapters did not rely entirely on student-based samples (with the exception of Pilot Study 1 and 4), they were nonetheless heavily reliant on the voluntary participation of University staff and students. In addition, a lack of consistency in priori power analyses across the studies and potential lack of power for some analyses further contribute to what is already a hot topic in Psychology (Maxwell, 2004; Bakker et al., 2012; Asendorpf et al., 2013). Some effort was made to mitigate sampling issues by following recommendations regarding the inclusion of effect sizes and, where possible, the confidence intervals of the effect sizes for the primary results (Maxwell, 2004). However, as Maxwell points out this is not a perfect solution and conducting priori power analyses as well as recruiting sufficiently large samples in order to detect potentially small effects should always be a priority. In terms of recommendations or lessons learned, there is nothing particularly novel to be said that has not already been covered by numerous articles (Maxwell, 2004; Bakker et al., 2012; Asendorpf et al., 2013). Given the nature of the work presented here (i.e., partial fulfilment of the requirements for a degree), Crutzen and Peters' (2017) stance on targeting student-led research in order to change the "norm" of conducting underpowered research is particularly relevant. However, mitigating the impact of time limitations and lack of resources which may affect the recruitment of large and diverse samples in student-led research is yet to be resolved.

In regards to the impact of the above mentioned issues, it would be fair to say that the weight of evidence a single study with sampling limitations provides is not sufficient on its own. Whether this is also true of a series of studies such as the ones presented in this thesis can be argued. On one hand, each study has its own sampling limitations with subsequent impact on statistical power and effect sizes. On the other hand, the results of the studies were obtained via different methods and largely support the importance of fiction in relation to people's attitudes toward, and perception of, robots which adds weight to the conclusions made so far. As mentioned earlier in this chapter, it would be interesting to see how and if the results of the studies in this thesis would be replicated with similar or different samples. Although - in line with the recommendations by Crutzen and Peters - a single large scale study with a more generalizable sample would be preferable over multiple smaller studies that may be underpowered and lead to biased conclusions. Finally, where future investigations relating to the impact of fictional representations of robots is concerned, estimates of the effect size should remain conservative (i.e., small effect) in order to ensure that the studies are sufficiently powered.

As touched upon in Section 6.1, only so much of the variability in people's attitudes can be explained by factors such as the robot's domain of application and individual differences between participants. Although this thesis takes into consideration a previously under-researched factor – namely, how individuals represent robots – it did not consider how representations may interact with other factors known to predict attitudes toward robots. For example, it is possible that one individual may have negative attitudes toward robots as companions because their representation centred around robots as emotionless machines, while another person may represent robots in a different way (e.g., as a helpful and friendly social agent) and thus feel differently about companion robots. Aa longitudinal study into how social and individual representations change over time and whether such changes lead to changes in attitudes may help to understand why there has been an overall decrease in people's positive attitudes toward robots in recent years (Gnambs & Appel, 2019).

Another area where further research is warranted, concerns the way in which representations and attitudes change as a function of demographic characteristics like age. Although the age of participants has previously been investigated as a factor that could account for some of the variability in people's attitudes, age differences with respect to differences in the fiction which people consume has yet to be investigated. The interviews carried out in Study 4 suggest that depictions of robots in fiction have changed over time and childhood experiences with such depictions may impact individuals' representation of robots. A comparison between the social representations of different age groups using the methodology used in Study 3 may be useful, as younger age groups have likely been exposed to different types of robots, in both fiction and real-life, as compared to older individuals. A formal mapping of how fictional (e.g., films and games) and non-fictional (e.g., newspaper articles) depictions of robots have changed over time may also be useful. Methods rooted in media studies, such as media content analysis (Macnamara, 2005), may be particularly helpful in establishing how the way robots are presented in media has changed over time and subsequently aid investigation into whether such changes are linked to changes in attitudes toward robots.

6.6 Conclusion

This thesis has shown that the concept of robots, and their role as a socially-relevant category, is not homogeneous across individuals. Not only do people imagine the appearance of robots in a variety of ways, but the very definition of what a robot is can be ambiguous and contradictory. Such contradictions appear to be, at least in part, the result of the gap between decades of fiction portraying robots as mostly socially inept, yet technologically advanced, and the somewhat less spectacular advancement of real robots which is only now catching up to fiction. If we are to understand how robots are perceived in their role as social agents, be it as companions or healthcare providers, we must consider how and why people's internal representation of what a robot is may clash with such roles.

References

Abric, J. C. (1993). Central system, peripheral system: their functions and roles in the dynamics of social representations. *Papers on Social Representations*, *2*, 75-78.

Age, U. K. (2017). Briefing: Health and Care of Older People in England 2017. *London: Age UK. Available at: www.ageuk.org.uk/Documents/EN-GB/For-professionals/Research/The_Health_and_Care_of_Older_People_in_England_2016.pdf*

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

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2014, October). 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 (ICRoM)* (pp. 754-759). IEEE. https://doi.org/10.1109/icrom.2014.6990994

Ando, H., Cousins, R., & Young, C. (2014). Achieving saturation in thematic analysis: Development and refinement of a codebook. *Comprehensive Psychology*, *3*, 03-CP. https://doi.org/10.2466/03.cp.3.4

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., ... & Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*(2), 108-119. https://doi.org/10.1002/per.1919

Avelino, J., Correia, F., Catarino, J., Ribeiro, P., Moreno, P., Bernardino, A., & Paiva, A. (2018, October). The power of a hand-shake in human-robot interactions. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1864-1869). IEEE. https://doi.org/10.1109/iros.2018.8593980

Backonja, U., Hall, A. K., Painter, I., Kneale, L., Lazar, A., Cakmak, M., ... & 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. https://doi.org/10.1111/jnu.12430

Bakker, M., Van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, *7*(6), 543-554. https://doi.org/10.1177/1745691612459060

Barnes, J., Richie, E., Lin, Q., Jeon, M., & Park, C. H. (2018). Emotive Voice Acceptance in Human-Robot Interaction. In *Proceedings of the 24th International Conference on Auditory Display*. https://doi.org/10.21785/icad2018.000

Bartneck, C., Suzuki, T., Kanda, T., & Nomura, T. (2007). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & Society*, *21*(1-2), 217-230. https://doi:10.1007/s00146-006-0052-7

Bazzano, F., & Lamberti, F. (2018). Human-robot interfaces for interactive receptionist systems and wayfinding applications. *Robotics*, *7*(3), 56. https://doi.org/10.3390/robotics7030056

Beasley, T. M., & Sheehan, J. K. (1994). *Choosing a MANOVA Test Statistic When Covariances Are Unequal* (ED379296). ERIC. https://files.eric.ed.gov/fulltext/ED379296.pdf

Bernotat, J., & Eyssel, F. (2018, August). 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)* (pp. 15-22). IEEE. https://doi.org/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. https://doi.org/10.1080/01449290802121230

Birnbaum, G. E., Mizrahi, M., Hoffman, G., Reis, H. T., Finkel, E. J., & Sass, O. (2016, March). Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 165-172). IEEE. https://doi.org/10.1109/hri.2016.7451748

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and*

Experiment, *2008*(10), 11-12. https://doi.org/10.1088/1742-5468/2008/10/p10008

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, *4*(1), 8-23. https://doi.org/10.1080/21670811.2015.1096598

Bradwell, H. L., Edwards, K. J., Winnington, R., Thill, S., & Jones, R. B. (2019). Companion robots for older people: importance of user-centred design demonstrated through observations and focus groups comparing preferences of older people and roboticists in South West England. *BMJ Open*, *9*(9), 1-13. https://doi.org/10.1136/bmjopen-2019-032468

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2012). *Thematic analysis.* In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA Handbooks in Psychology*. *APA Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, And Biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-000

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. https://doi.org/10.1037/0022-3514.47.6.1191

Broadbent, E., Kerse, N., Peri, K., Robinson, H., Jayawardena, C., Kuo, T., ... & 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. https://doi.org/10.1111/ajag.12190

Broadbent, E., Kuo, I. H., Lee, Y. I., Rabindran, J., Kerse, N., Stafford, R., & MacDonald, B. A. (2010). Attitudes and reactions to a healthcare robot. *Telemedicine and e-Health*, *16*(5), 608-613. https://doi:10.1089/tmj.2009.0171

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. https://doi.org/10.1007/s12369-009-0030-6

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. https://doi.org/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. https://doi.org/10.4017/gt.2009.08.02.002.00

Bruckenberger, U., Weiss, A., Mirnig, N., Strasser, E., Stadler, S., & Tscheligi, M. (2013, October). The good, the bad, the weird: audience evaluation of a "Real" robot in relation to science fiction and mass media. In *International Conference on Social Robotics* (pp. 301-310). Springer, Cham. https://doi.org/10.1007/978-3-319-02675-6_30

Bumby, K., & Dautenhahn, K. (1999, August). Investigating children's attitudes towards robots: A case study. In *Proc. CT99, The Third International Cognitive Technology Conference* (pp. 391-410). https://doi.org/10.1109/humanoids.2014.7041385

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*(2), 255-280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or not to Bonferroni: when and how are the questions. *Bulletin of the Ecological Society of America*, *81*(3), 246-248.

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. https://doi.org/10.1080/09540091.2018.1454889

Carrillo, F. M., Butchart, J., Kruse, N., Scheinberg, A., Wise, L., & McCarthy, C. (2018, August). 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) (pp. 850-855). IEEE. https://doi.org/10.1109/roman.2018.8525508

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

Cavallo, E., Galiani, S., Noy, I., & Pantano, J. (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, *95*(5), 1549-1561. https://doi.org/10.1109/roman.2014.6926258

Cavallo, F., Esposito, R., Limosani, R., Manzi, A., Bevilacqua, R., Felici, E., ... & Dario, P. (2018). Robotic services acceptance in smart environments with older adults: user satisfaction and acceptability study. *Journal of Medical Internet Research*, *20*(9), e264. https://doi.org/10.2196/jmir.9460

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. https://doi.org/10.1016/j.chb.2018.02.025

Chang, W. L., White, J. P., Park, J., Holm, A., & Šabanović, S. (2012, September). 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* (pp. 845-850). IEEE. https://doi.org/10.1109/roman.2012.6343857

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

Cho, H., Shen, L., & Wilson, K. (2014). Perceived realism: Dimensions and roles in narrative persuasion. *Communication Research*, *41*(6), 828-851. https://doi:10.1177/0093650212450585

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661-703. https://doi.org/10.1137/070710111

Coco, K., Kangasniemi, M., & Rantanen, T. (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. https://doi.org/10.1111/jnu.12435

Conti, D., Cattani, A., Di Nuovo, S., & Di Nuovo, A. (2015, August). 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)* (pp. 555-560). IEEE. https://doi.org/10.1109/roman.2015.7333601

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. https://doi.org/10.1007/s12369-016-0359-6

Correia, F., Guerra, C., Mascarenhas, S., Melo, F. S., & Paiva, A. (2018, July). Exploring the impact of fault justification in human-robot trust. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems* (pp. 507-513). https://doi.org/10.1007/s12369-016-0359-6

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. https://doi.org/10.1002/cav.317

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. https://doi.org/10.1177/0146167294206001

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2003). Efficiency of scale-free networks: error and attack tolerance. *Physica A: Statistical Mechanics and its Applications*, *320*, 622-642. https://doi.org/10.1016/s0378-4371(02)01545-5

Crutzen, R., & Peters, G. J. Y. (2017). Targeting next generations to change the common practice of underpowered research. *Frontiers in Psychology*, 8, 1-4. https://doi.org/10.3389/fpsyg.2017.01184

Cuneo, J. (2011). “Hello, Computer”: The Interplay of Star Trek and Modern Computing. *Science Fiction and Computing: Essays on Interlinked Domains, eds. David L. Ferroand Eric G. Swedin*, 131-147.

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), 1-13. https://doi.org/10.3389/fpsyg.2015.01701

Dany, L., Urdapilleta, I., & Monaco, G. L. (2015). Free associations and social representations: some reflections on rank-frequency and importance-frequency methods. *Quality & Quantity*, *49*(2), 489-507. https://doi.org/10.1007/s11135-014-0005-z

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340. https://doi.org/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). https://10.1136/bmjpo-2018-000371

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

de Graaf, M. M., & Allouch, S. B. (2013, August). The relation between people's attitude and anxiety towards robots in human-robot interaction. In *2013 IEEE RO-MAN* (pp. 632-637). IEEE.

de Graaf, M. M., Allouch, S. B., & Lutfi, S. (2016, August). 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)* (pp. 1077-1083). IEEE. https://doi.org/10.1109/roman.2013.6628419

de Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. (2000). The mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, *50*(1), 1-18. https://doi.org/10.1016/S0169-7439(99)00047-7 de Rosa, A. (1993). Social representations and attitudes: Problems of coherence between the theoretical definition and procedure of research. *Papers on social representations*, *2*, 178-192. doi.org/10.4324/9780203102138

Dinet, J., & Vivian, R. (2014). Exploratory investigation of attitudes towards assistive robots for future users. *Le travail humain*, *77*(2), 105-125. https://doi.org/10.3917/th.772.0105

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002, June). All robots are not created equal: the design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 321-326). ACM. https://doi.org/10.1145/778712.778756

Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., ... & Zhang, H. (2011, August). Understanding the impact of video quality on user engagement. In *ACM SIGCOMM Computer Communication Review* (Vol. 41, No. 4, pp. 362-373). ACM. https://doi.org/10.1145/2043164.2018478

Doerfel, M. L. (1998). What constitutes semantic network analysis? A comparison of research and methodologies. *Connections*, *21*(2), 16-26. https://doi.org/10.1111/j.1468-2958.1999.tb00463.x

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3-4), 177-190. https://doi.org/10.1016/s0921-8890(02)00374-3

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*.

Ebel, K. H. (1986). The impact of industrial robots on the world of work. *International Labour Review*, *125*, 65-72. https://doi.org/10.1016/0167-8493(87)90034-9

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. https://10.1007/s12369-011-0094-y

Erebak, S., & Turgut, T. (2019). Caregivers' attitudes toward potential robot coworkers in elder care. *Cognition, Technology & Work*, *21*(2), 327-336. https://doi.org/10.1007/s10111-018-0512-0

European Commission (2012, September). *Public Attitudes towards Robots.* Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_382_en.pdf

Eyssel, F., De Ruiter, L., Kuchenbrandt, D., Bobinger, S., & Hegel, F. (2012, March). '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)* (pp. 125-126). IEEE. https://doi.org/10.1145/2157689.2157717

Ezer, N., Fisk, A. D., & Rogers, W. A. (2009, July). Attitudinal and intentional acceptance of domestic robots by younger and older adults. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 39-48). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02710-9_5

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*(1), 74-147. https://doi.org/10.1080/10463280600681248

Finch, H. (2005). Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology*, *1*(1), 27-38. https://10.1027/1614-1881.1.1.27

Fink, J. (2012, October). Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *International Conference on Social Robotics* (pp. 199-208). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34103-8_20

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. https://doi.org/10.1515/pjbr-2018-0007

French, A., Macedo, M., Poulsen, J., Waterson, T., & Yu, A. (2008). Multivariate analysis of variance (MANOVA). *San Francisco State University*. https://doi.org/10.1007/springerreference_205476

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

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2 - 18. https://doi.org/10.1037/a0024338

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. https://10.1016/j.chb.2016.03.057

Giannakouris, K. (2010). Regional population projections EUROPOP2008: Most EU regions face older population profile in 2030. *Eurostat: Statistics in Focus*, *1*(2010), 1-19.

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. https://doi.org/10.1016/j.chb.2018.11.045

Gockley, R., Forlizzi, J., & Simmons, R. (2007, March). Natural person-following behavior for social robots. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on* (pp. 17-24). IEEE. https://doi.org/10.1145/1228716.1228720

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. https://doi.org/10.1177/0278364918778344

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, *38*(2), 247-266. https://10.1207/s15326950dp3802_5

Green, M. C., Garst, J., & Brock, T. C. (2004). *The power of fiction: Determinants and boundaries. The psychology of entertainment media: Blurring the lines between entertainment and persuasion*, 161-176. Lawrence Erlbaum Associates.

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. https://doi.org/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. https://doi.org/10.1037/0022-3514.85.2.197

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. https://doi.org/10.1111/1468-0009.00223

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

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

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. https://doi.org/10.1109/roman.2015.7333613

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. https://doi.org/10.1109/kst.2016.7440504

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. https://doi.org/10.1007/s12369-010-0068-5

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

Herse, S., Vitale, J., Tonkin, M., Ebrahimian, D., Ojha, S., Johnston, B., ... & Williams, M. A. (2018, August). 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. https://doi.org/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. https://doi.org/10.1111/j.2044-8309.2011.02047.x

Hockstein, N. G., Gourin, C. G., Faust, R. A., & Terris, D. J. (2007). A history of robots: from science fiction to surgical robotics. *Journal of Robotic Surgery*, *1*(2), 113-118. https://doi.org/10.1007/s11701-007-0021-2

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. https://doi.org/10.1109/roman.2017.8172294

Howarth, C. (2006). How social representations of attitudes have informed attitude theories: the consensual and the reified. *Theory & Psychology*, *16*(5), 691-714. https://doi.org/10.1177/0959354306067443

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Idoiaga Mondragon, N., Gil de Montes, L., & Valencia, J. (2017). Understanding an Ebola outbreak: Social representations of emerging infectious diseases. *Journal* *of Health Psychology*, *22*(7), 951-960. https://doi.org/10.1177/1359105315620294

Iwamura, Y., Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2011, March). 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* (pp. 449-456). https://doi.org/10.1145/1957656.1957816

Jayawardena, C., Kuo, I., Datta, C., Stafford, R. Q., Broadbent, E., & MacDonald, B. A. (2012, June). Design, implementation and field tests of a socially assistive robot for the elderly: HealthBot Version 2. In *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on* (pp. 1837-1842). IEEE. https://doi.org/10.1109/biorob.2012.6290890

Johnson, D. R., Jasper, D. M., Griffin, S., & Huffman, B. L. (2013). Reading narrative fiction reduces Arab-Muslim prejudice and offers a safe haven from intergroup anxiety. *Social Cognition*, *31*(5), 578-598. https://doi.org/10.1521/soco.2013.31.5.578

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. https://doi.org/10.1080/10447318.2013.873278

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Pychology*, *91*(1), 16. https://doi.org/10.1037/0022-3514.91.1.16

Keczer, Z., File, B., Orosz, G., & Zimbardo, P. G. (2016). Social representations of hero and everyday hero: A network study from representative samples. *PLOS ONE*, *11*(8), 1-17. https://doi.org/10.1371/journal.pone.0159354

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. https://doi.org/10.3390/fi10100101

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. https://doi.org/10.1016/j.chb.2013.02.009

Kim, M. G., Lee, J., Aichi, Y., Morishita, H., & Makino, M. (2016, November). 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. https://doi.org/10.1109/mhs.2016.7824236

Klamer, T., & Allouch, S. B. (2010, March). 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. https://doi.org/10.4108/icst.pervasivehealth2010.8892

Kraus, M., Kraus, J., Baumann, M., & Minker, W. (2018, May). 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)*. https://doi.org/10.4018/ijmhci.2019070101

Kriz, S., Ferro, T. D., Damera, P., & Porter, J. R. (2010, September). 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. https://doi.org/10.1109/ROMAN.2010.5598620

Kuchenbrandt, D., & Eyssel, F. (2012, September). 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. https://doi.org/10.1109/roman.2012.6343795

Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R. M. Q., & MacDonald, B. A. (2009, September). 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. https://doi.org/10.1109/roman.2009.5326292

Lahlou, S., & Abric, J. C. (2011). What are the" elements" of a representation? *Papers on social representations*, *20*(2), 20-1.

Lamarque, P., & Olsen, S. (1996). Setting the Scene. In Truth, Fiction, and Literature: A Philosophical Perspective. : Oxford University Press. Retrieved 27 Jan. 2020, from https://www-oxfordscholarship-com.sheffield.idm.oclc.org/view/10.1093/acprof:oso/9780198236818.001.0001/acprof-9780198236818-chapter-1.

Lehmann, C., Pellicciari, M., Drust, M., & Gunnink, J. W. (2013, June). Machining with industrial robots: the COMET project approach. In *International Workshop on Robotics in Smart Manufacturing* (pp. 27-36). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39223-8_3

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

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. https://doi.org/10.1007/s12369-010-0056-9

Li, L., Alderson, D., Doyle, J. C., & Willinger, W. (2005). Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics*, *2*(4), 431-523. https://doi.org/10.1080/15427951.2005.10129111

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. https://doi.org/10.1089/cyber.2016.0199

Lord, C. G., & Lepper, M. R. (1999). Attitude representation theory. *Advances in Experimental Social Psychology*, *31*, 265-343. https://doi.org/10.1016/S0065-2601(08)60275-0

Lorenčík, D., Tarhaničová, M., & Sinčák, P. (2013, January). Influence of sci-fi films on artificial intelligence and vice-versa. In *Applied Machine Intelligence and Informatics (SAMI), 2013 IEEE 11th International Symposium on* (pp. 27-31). IEEE. https://doi.org/10.1109/SAMI.2013.6480990

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. https://doi.org/10.1080/10400435.2013.869703

Lugrin, B., Dippold, J., & Bergmann, K. (2018, October). 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. https://doi.org/10.1109/iros.2018.8593818

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. https://doi.org/10.1007/s00146-008-0181-2

Macnamara, J. R. (2005). Media content analysis: Its uses, benefits and best practice methodology. *Asia Pacific Public Relations Journal*, *6*(1), 1-34.

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. https://doi.org/10.1109/roman.2018.8525796

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. https://doi.org/10.1007/s00146-011-0370-2

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: causes, consequences, and remedies. *Psychological Methods*, 9(2), 147-163. https://doi.org/10.1037/1082-989X.9.2.147

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. https://doi.org/10.1007/s12369-016-0390-7

McKitterick. W. (2016) The global smartphone report: The forces behind the global deceleration in smartphone sales. Bus 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

Mehta, C. R., & Patel, N. R. (2011). IBM SPSS exact tests. *Armonk, NY: IBM Corporation*.

Meissner, F., Grigutsch, L. A., Koranyi, N., Müller, F., & Rothermund, K. (2019). Predicting behavior with implicit measures: Disillusioning findings, reasonable explanations, and sophisticated solutions. *Frontiers in Psychology*, *10*, 1-16. https://doi.org/10.3389/fpsyg.2019.02483

Mendelson, A. L., & Papacharissi, Z. (2007). Reality vs. Fiction how defined realness affects cognitive & emotional responses to photographs. *Visual Communication Quarterly*, *14*(4), 231-243. https://doi.org/10.1080/15551390701730224

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*.

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. https://doi.org/10.1109/roman.2016.7745106

Mubin, O., Obaid, M., Sandoval, E., & Fjeld, M. (2015, October). Using video preferences to understand the human perception of real and fictional robots. In *Proceedings of the 3rd International Conference on Human-Agent Interaction* (pp. 33-39). ACM. https://doi.org/10.1145/2814940.2814958

Mulligan, K., & Habel, P. (2011). An experimental test of the effects of fictional framing on attitudes. *Social Science Quarterly*, *92*(1), 79-99. https://doi.org/10.1111/j.1540-6237.2011.00758.x

Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). Conducting semi-structured interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation* (4th ed., pp. 492-505). Jossey-Bass. https://doi.org/10.1002/9781119171386.ch19

Nomura, T. (2017, August). 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. https://doi.org/10.1109/roman.2017.8172354

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

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2004, September). 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. https://doi.org/10.1109/roman.2004.1374726

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2005, August). People's assumptions about robots: Investigation of their relationships with attitudes and emotions toward robots. In *IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2005.* (pp. 125-130). IEEE. https://doi.org/10.1109/ROMAN.2005.1513768

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. https://doi.org/10.1109/TRO.2007.914004

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2009). Age differences and images of robots: Social survey in Japan. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 10*(3), 374-391. https://doi.org/10.1075/is.10.3.05nom

Nomura, T., Shintani, T., Fujii, K., & Hokabe, K. (2007, March). 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* (pp. 13-18). https://doi.org/doi.org/10.5772/4692

Nomura, T., Shintani, T., Fujii, K., & Hokabe, K. (2007, March). 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* (pp. 13-18). https://doi.org/10.5772/4692

Nørskov, M. (2017). Technological dangers and the potential of human–robot interaction: a philosophical investigation of fundamental epistemological mechanisms of discrimination. In *Social Robots* (pp. 99-121). Routledge. https://doi.org/10.4324/9781315563084-6

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. *Automatic Processes in Social Thinking and Behavior*, *4*, 265-292.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology*, *25*(3), 241-286. https://doi.org/10.1006/ceps.2000.1040

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

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology*, *108*(4), 562–571. https://doi.org/10.1037/pspa0000023

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814-818. https://doi.org/10.1038/nature03607

Park, E., & Del Pobil, A. P. (2012). An Acceptance Model for Service Robots in Global Markets. *International Journal of Humanoid Robotics*, *9*(04), 1250026. https://doi.org/10.1142/s0219843612500260

Park, E., & Kwon, S. J. (2016). The adoption of teaching assistant robots: a technology acceptance model approach. *Program*, *50*(4), 354-366. https://doi.org/10.1108/prog-02-2016-0017

Park, E., & Lee, J. (2014). I am a warm robot: the effects of temperature in physical human-robot interaction. *Robotica*, *32*(1), 133-142. https://doi.org/10.1017/s026357471300074x

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65-85. https://doi.org/10.1146/annurev.psych.49.1.65

Pettit, H. (2018). Death of 75 million jobs worldwide: Robots and artificial intelligence will take over HALF of all tasks in the workplace by 2025, warns World Economic Forum. *The Daily Mail Online*. Retrieved from https://www.dailymail.co.uk/sciencetech/article-6175599/Machines-handle-half-workplace-tasks-2025.html

Piçarra, N., Giger, J. C., Pochwatko, G., & Gonçalves, G. (2016). Making sense of social robots: A structural analysis of the layperson's social representation of robots. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *66*(6), 277-289. https://doi.org/10.1016/j.erap.2016.07.001

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.

Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis. *The Annals of Mathematical Statistics*, *26*(1), 117-121. https://doi.org/10.1214/aoms/1177728599

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*, 141-141. https://doi.org/10.3389/fnagi.2015.00141

Pouliot, L., & Cowen, P. S. (2007). Does perceived realism really matter in media effects?. *Media Psychology*, *9*(2), 241-259. https://doi.org/10.1080/15213260701285819

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. https://doi.org/10.1080/10463280601055772

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

Prescott, T. J., et al. (2012, May). Robot Companions For Citizens: Roadmapping the potential for future robots in empowering older people. In *Bridging Research in Ageing and ICT Development Conference*, Prague, Czech Republic.

Randelli, G., Venanzi, M., & Nardi, D. (2011, July). Evaluating tangible paradigms for ground robot teleoperation. In *2011 RO-MAN* (pp. 389-394). IEEE. https://doi.org/10.1109/roman.2011.6005240

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. https://doi.org/10.1111/scs.12508

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

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. https://doi.org/10.1007/s12369-015-0308-9

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. https://doi.org/10.1016/j.chb.2018.09.041

Riek, L. D., Adams, A., & Robinson, P. (2011, March). Exposure to cinematic depictions of robots and attitudes towards them. In *Proceedings of International Conference on Human-Robot Interaction, Workshop on Expectations and Intuitive Human-Robot Interaction*. https://doi.org/10.1.1.442.5427

Robert, D., & van den Bergh, V. (2014, August). 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. https://doi.org/10.1109/roman.2014.6926372

Robillard, J. M., & Kabacińska, K. (2020). Realizing the potential of robotics for aged care through co-creation. *Journal of Alzheimer's Disease*, *76*, 461-466. https://doi.org/10.3233/JAD-200214

Robot-Era Project [RobotEraProject]. (2014). *Robot-Era Project Official Video* [Video File]. Available from https://www.youtube.com/watch?v=lv43z8YVQkY&t=1s

Robot-Era Project [Stanislava Naneva]. (2017). *Robot-Era Project 640 x 480* [Video File]. Available from https://www.dailymotion.com/video/x6wv4w1

Rosén, J., Richardson, K., Lindblom, J., & Billing, E. (2018). The robot illusion: facts and fiction. In *Workshop in Explainable Robotics System, workshop in conjunction with HRI Human Robot Interaction 2018, Chicago, USA, March, 5-8, 2018*.

Rosenthal-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. https://doi.org/10.1016/j.chb.2014.03.066

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. https://doi.org/10.1007/s12369-013-0196-9

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015, March). 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. https://doi.org/10.1145/2696454.2696497

Salmons, J. (2014). Chapter 10: Conducting the online interview. In V. Knight (Ed.), *Qualitative Online Interviews: Strategies, Design, and Skills* (2nd ed., pp. 194-219). Sage Publications.

Salvini, P., Laschi, C., & Dario, P. (2010). Design for acceptability: improving robots' coexistence in human society. *International Journal of Social Robotics*, *2*(4), 451-460. https://doi.org/10.1007/s12369-010-0079-2

Sandoval, E. B., Mubin, O., & Obaid, M. (2014, October). Human robot interaction and fiction: A contradiction. In *International Conference on Social Robotics* (pp. 54-63). Springer, Cham. https://doi.org/10.1007/978-3-319-11973-1_6

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. https://doi.org/10.1007/s12369-017-0452-5

Schreier, J. (2012). Robot & Frank. *Science fiction comedy-drama film, directed by Jake Schreier. Samuel Goldwyn Films*.

Schwartz, H. A., & Ungar, L. H. (2015). Data-driven content analysis of social media: a systematic overview of automated methods. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 78-94. https://doi.org/10.1177/0002716215569197

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. https://doi.org/10.1007/s10209-005-0118-1

Serholt, S., Basedow, C. A., Barendregt, W., & Obaid, M. (2014, November). 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. https://doi.org/10.1109/humanoids.2014.7041511

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. https://doi.org/10.1075/is.12.3.04shi

Smarr, C. A., Prakash, A., Beer, J. M., Mitzner, T. L., Kemp, C. C., & Rogers, W. A. (2012, September). Older adults' preferences for and acceptance of robot assistance for everyday living tasks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 153-157). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/1071181312561009

Smith, N., & Joffe, H. (2013). How the public engages with global warming: A social representations approach. *Public Understanding of Science*, *22*(1), 16-32. https://doi.org/10.1177/0963662512440913

Spence, P. R., Edwards, A., & Edwards, C. (2018, March). 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* (pp. 243-244). https://doi.org/10.1145/3173386.3177071

Stadler, S., Weiss, A., & Tscheligi, M. (2014, August). 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. https://doi.org/10.1109/roman.2014.6926388

Stafford, R. Q., Broadbent, E., Jayawardena, C., Unger, U., Kuo, I. H., Igic, A., ... & MacDonald, B. A. (2010, September). 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. https://doi.org/10.1109/roman.2010.5598679

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. https://doi.org/10.1007/s12369-017-0422-y

Steyn Jr, H. S., & Ellis, S. M. (2009). Estimating an effect size in one-way multivariate analysis of variance (MANOVA). *Multivariate Behavioral Research*, *44*(1), 106-129. https://doi.org/10.1080/00273170802620238

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41-78. https://doi.org/10.1207/s15516709cog2901_3

Stuck, R. E., & Rogers, W. A. (2018). Older adults' perceptions of supporting factors of trust in a robot care provider. *Journal of Robotics*, *2018*.

Sundar, S. S., Waddell, T. F., & Jung, E. H. (2016, March). 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. https://doi.org/10.1109/hri.2016.7451771

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. https://doi.org/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. In *Proceedings of Adaptive and Emergent Behaviour and Complex Systems* (pp. 109-11). SSAISB. https://doi.org/10.1109/achi.2009.62

Takayama, L., Ju, W., & Nass, C. (2008, March). Beyond dirty, dangerous and dull: what everyday people think robots should do. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (pp. 25-32). ACM. https://doi.org/10.1145/1349822.1349827

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. https://doi.org/10.1109/tie.2007.891642

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. https://doi.org/10.1016/j.chb.2014.05.014

Thellman, S., & Ziemke, T. (2017, March). Social attitudes toward robots are easily manipulated. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 299-300). ACM. https://doi.org/10.1145/3029798.3038336

Todorov, V., & Filzmoser, P. (2010). Robust Statistic for the One-way MANOVA. *Computational Statistics & Data Analysis*, *54*(1), 37-48. https://doi.org/10.1016/j.csda.2009.08.015

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. https://doi.org/10.1007/s10846-013-0019-0

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

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

van den 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*, *6*(4), 519-531. https://doi.org/10.1007/s12369-014-0231-5

van den Bulck, J. J. (2002). The impact of television fiction on public expectations of survival following inhospital cardiopulmonary resuscitation by medical professionals. *European Journal of Emergency Medicine*, *9*(4), 325-329. https://doi.org/10.1097/01.emj.0000043720.82722.66

Vanman, E. J., & Kappas, A. (2019). "Danger, Will Robinson!" The challenges of social robots for intergroup relations. *Social and Personality Psychology Compass*, *13*(8), e12489. https://doi.org/10.1111/spc3.12489

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

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. https://doi.org/10.1177/0018720818811190

Wachelke, J. F. R., & Lins, S. (2008). Changing masks: a making effect on young people's social representation on aging?

Wagner, W., Valencia, J., & Elejabarrieta, F. (1996). Relevance, discourse and the 'hot'stable core social representations—A structural analysis of word associations. *British Journal of Social Psychology*, *35*(3), 331-351. https://doi.org/10.1111/j.2044-8309.1996.tb01101.x

Warren, R., Smith, R. F., & Cybenko, A. K. (2011). *Use of Mahalanobis distance for detecting outliers and outlier clusters in markedly non-normal data: a vehicular traffic example*. SRA INTERNATIONAL INC DAYTON OH.

Weiss, A., Bernhaupt, R., Tscheligi, M., & Yoshida, E. (2009, April). 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 (Edinburgh, 8-9 April 2009), SSAISB* (pp. 150-157).

Weiss, A., Bernhaupt, R., Tscheligi, M., Wollherr, D., Kuhnlenz, K., & Buss, M. (2008, August). 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. https://doi.org/10.1109/roman.2008.4600751

Wissmath, B., Weibel, D., & Groner, R. (2009). Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment. *Journal of Media Psychology*, *21*(3), 114-125. https://doi.org/10.1027/1864-1105.21.3.114

Woods, S., Walters, M., Koay, K. L., & Dautenhahn, K. (2006, March). Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *9th IEEE International Workshop on Advanced Motion Control, 2006.* (pp. 750-755). IEEE. https://doi.org/10.1109/AMC.2006.1631754

Wu, Y. H., Fassert, C., & Rigaud, A. S. (2012). Designing robots for the elderly: appearance issue and beyond. *Archives of Gerontology and Geriatrics*, *54*(1), 121-126. https://doi.org/10.1016/j.archger.2011.02.003

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. https://doi.org/10.2147/cia.s56435

Wullenkord, R., & Eyssel, F. (2014, August). Improving attitudes towards social robots using imagined contact. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 489-494). IEEE. https://doi.org/10.1109/roman.2014.6926300

Wullenkord, R., Fraune, M. R., Eyssel, F., & Šabanović, S. (2016, August). 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. https://doi.org/10.1109/roman.2016.7745228

Zhang, T., Zhu, B., Lee, L., & Kaber, D. (2008, August). Service robot anthropomorphism and interface design for emotion in human-robot interaction.

In *2008 IEEE International Conference on Automation Science and Engineering* (pp. 674-679). IEEE. https://doi.org/10.1109/COASE.2008.4626532

Zhao, D., & Yi, J. (2006, September). Robot planning with artificial potential field guided ant colony optimization algorithm. In *International Conference on Natural Computation* (pp. 222-231). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11881223_28

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. https://doi.org/10.3389/fpsyg.2015.00883

Appendix A

Checklist used to operationalise the definition of a social robot in Chapter 2.

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).

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).

Appendix B

Tool used to assess the methodological quality of studies included in Chapter 2.

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* ≥ 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* ≥ 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* ≥ 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* ≥ 0.7).
4.0	Test-retest reliability was measured and reported adequately within the study, the measure's reliability was excellent (correlation of *r* ≥ 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 > α); 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 > α ≥ 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.

Appendix C

Graphical representation of the factors that influence the main outcomes in Chapter 2.

Type of exposure to robots

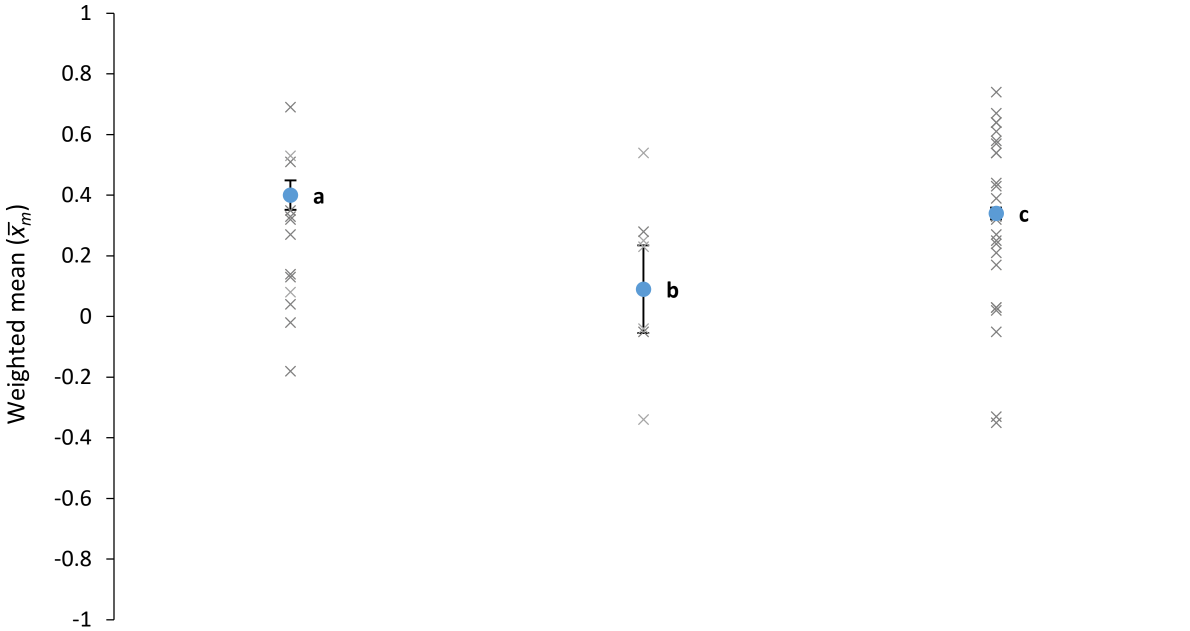


Fig 1. Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

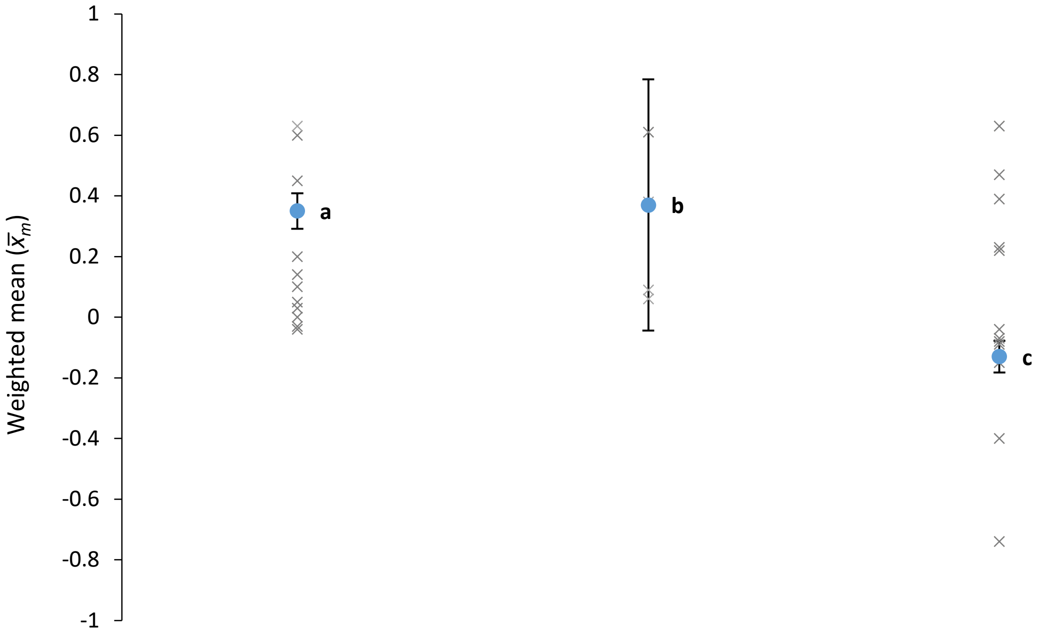


Fig 2. Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

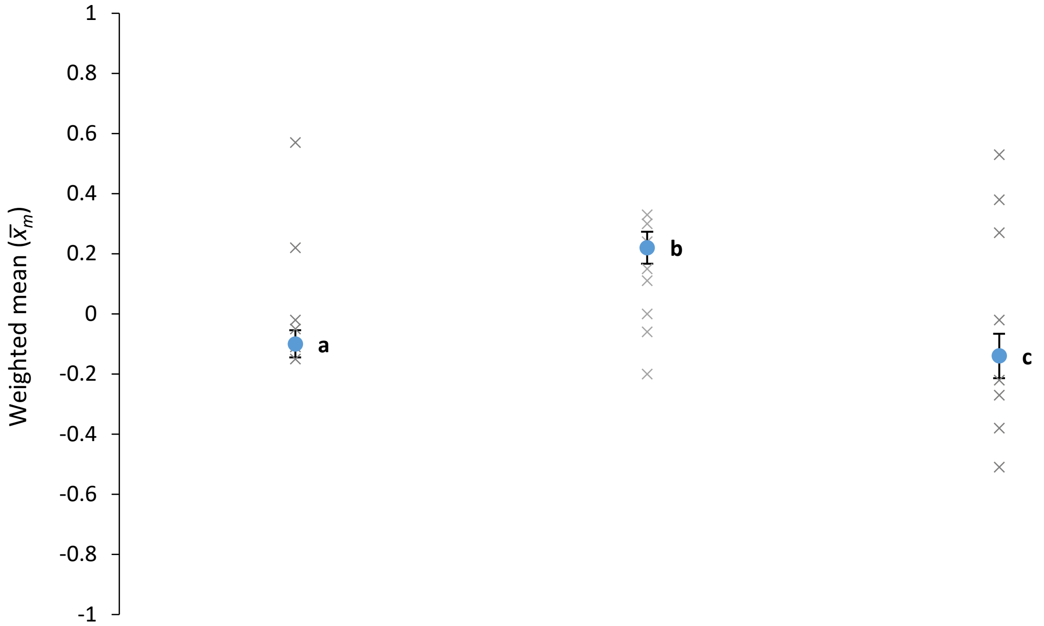


Fig 3. General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

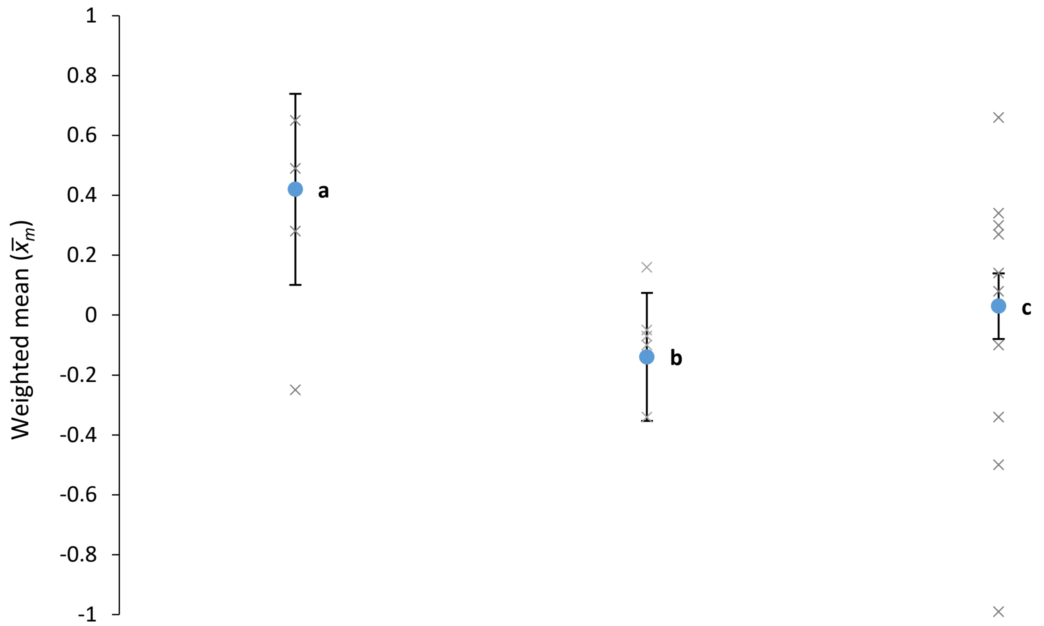


Fig 4. Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

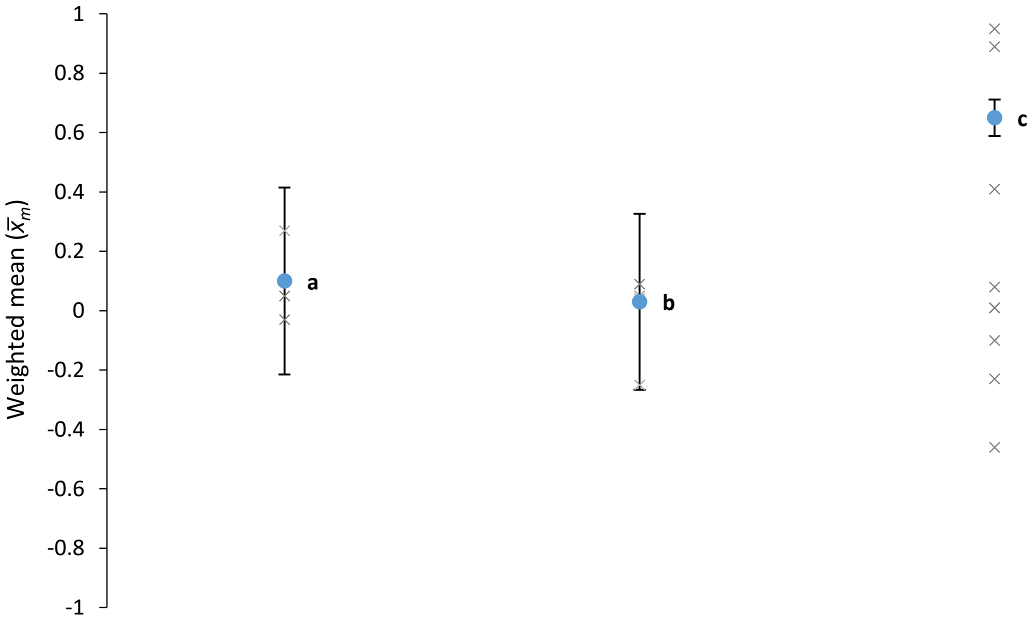


Fig 5. Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

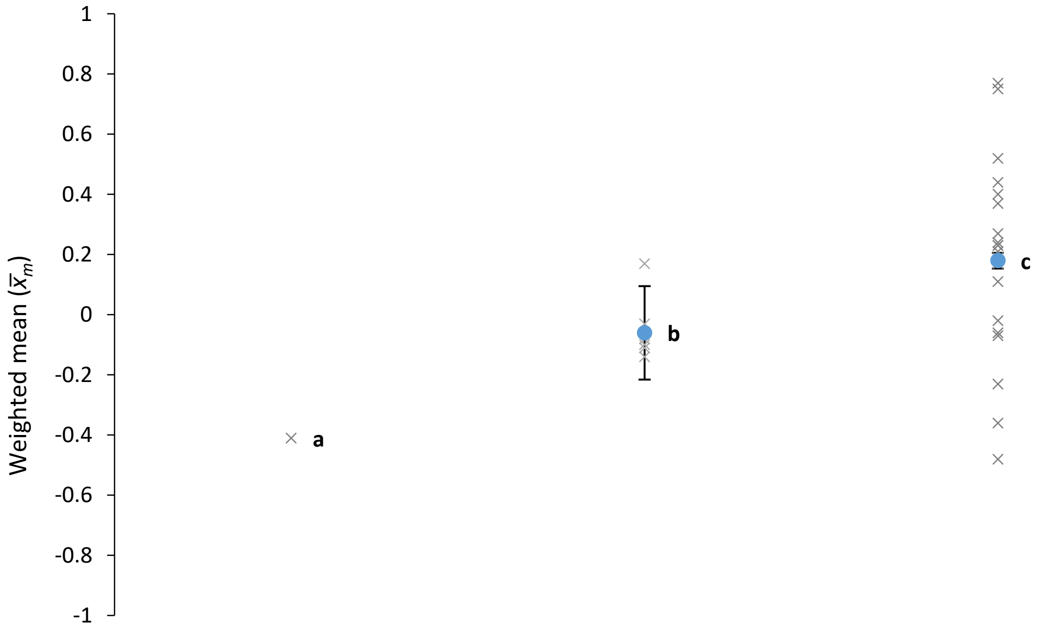


Fig 6. Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of exposure to robots, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) No HRI group; b) Indirect HRI group; and c) Direct HRI group.

Domain of application

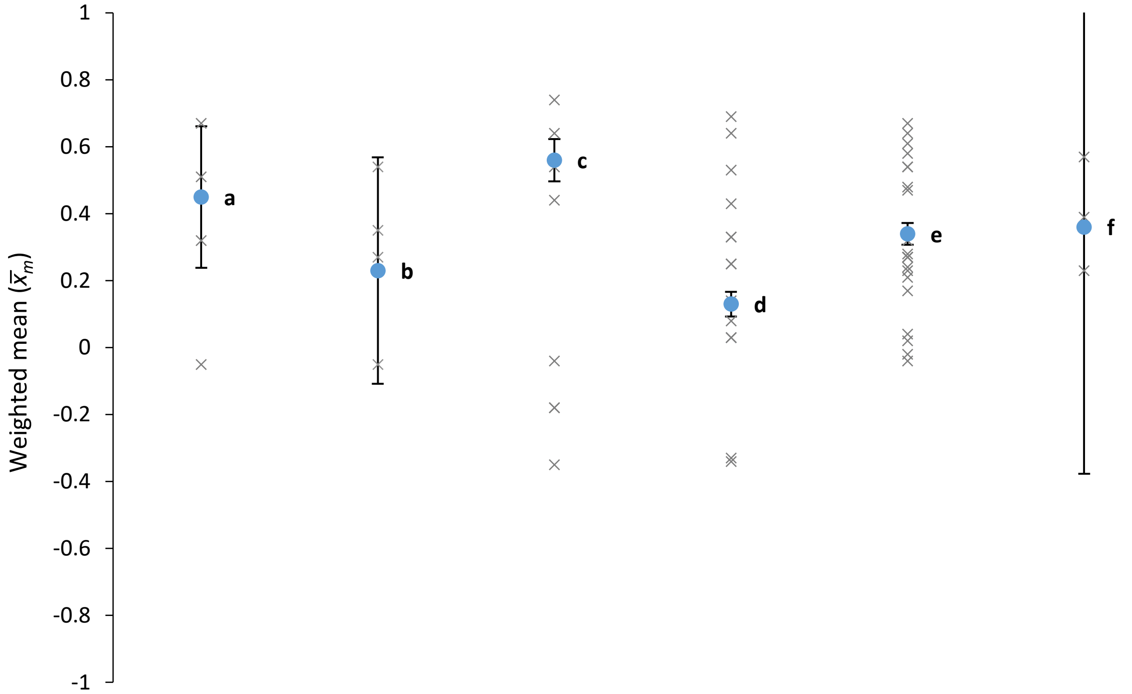


Fig 7. Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each 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.

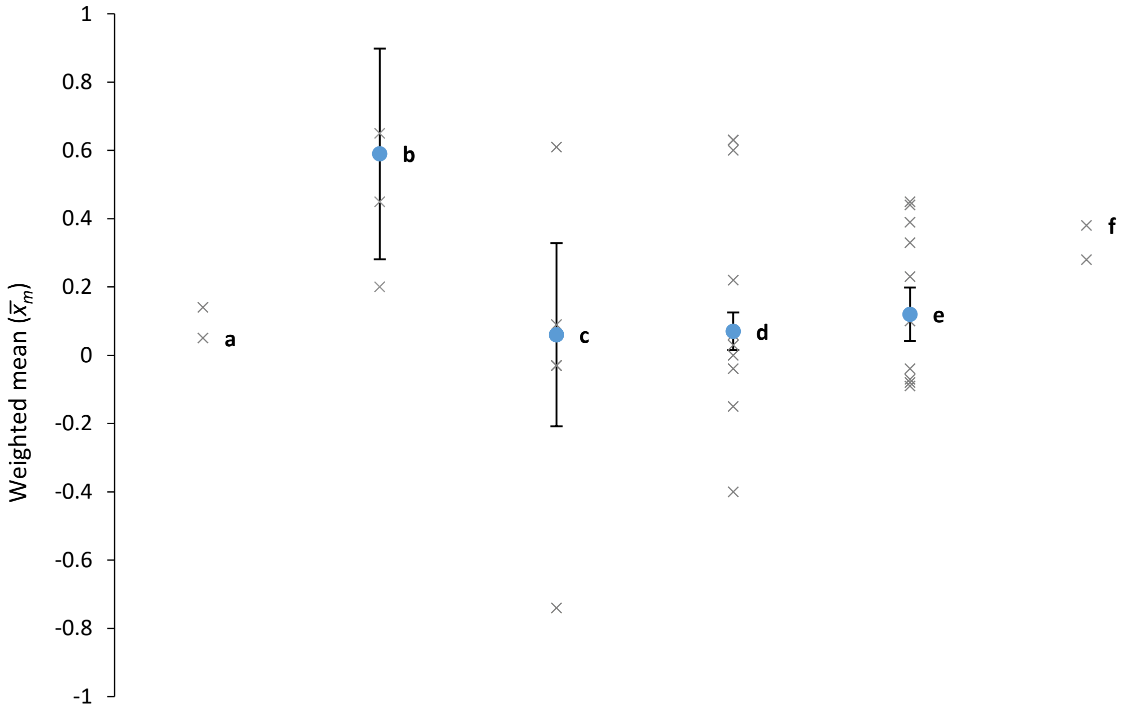


Fig 8. Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each 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.

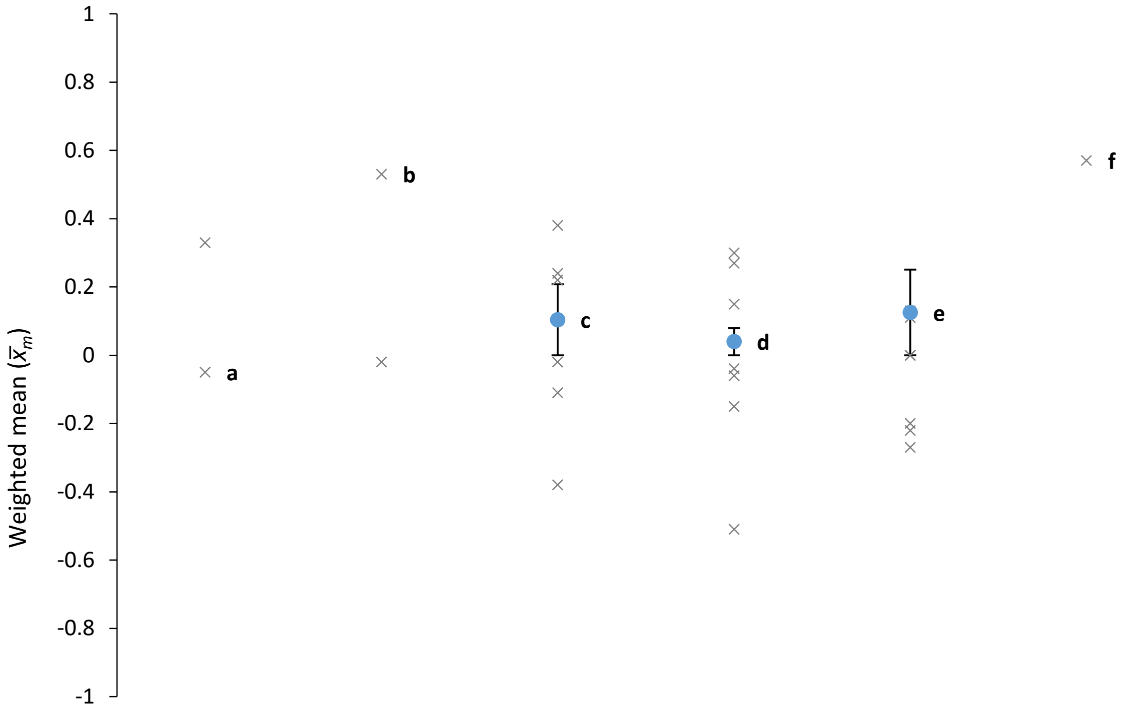


Fig 9. General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each 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.

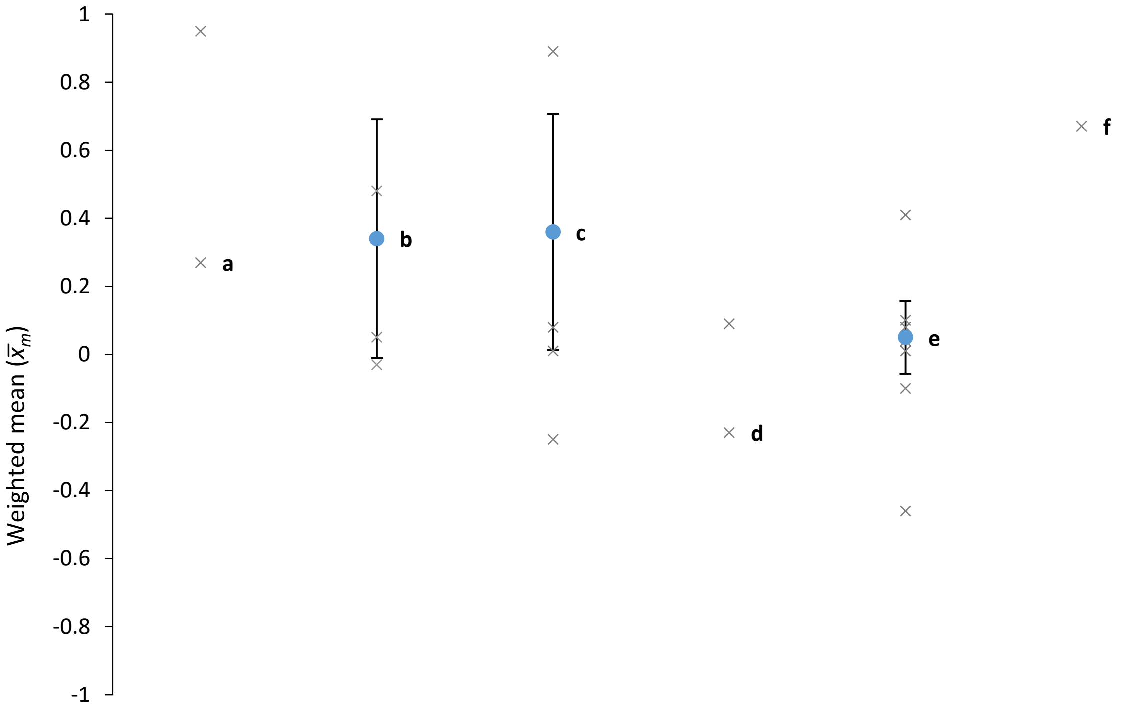


Fig 11. Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in 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.

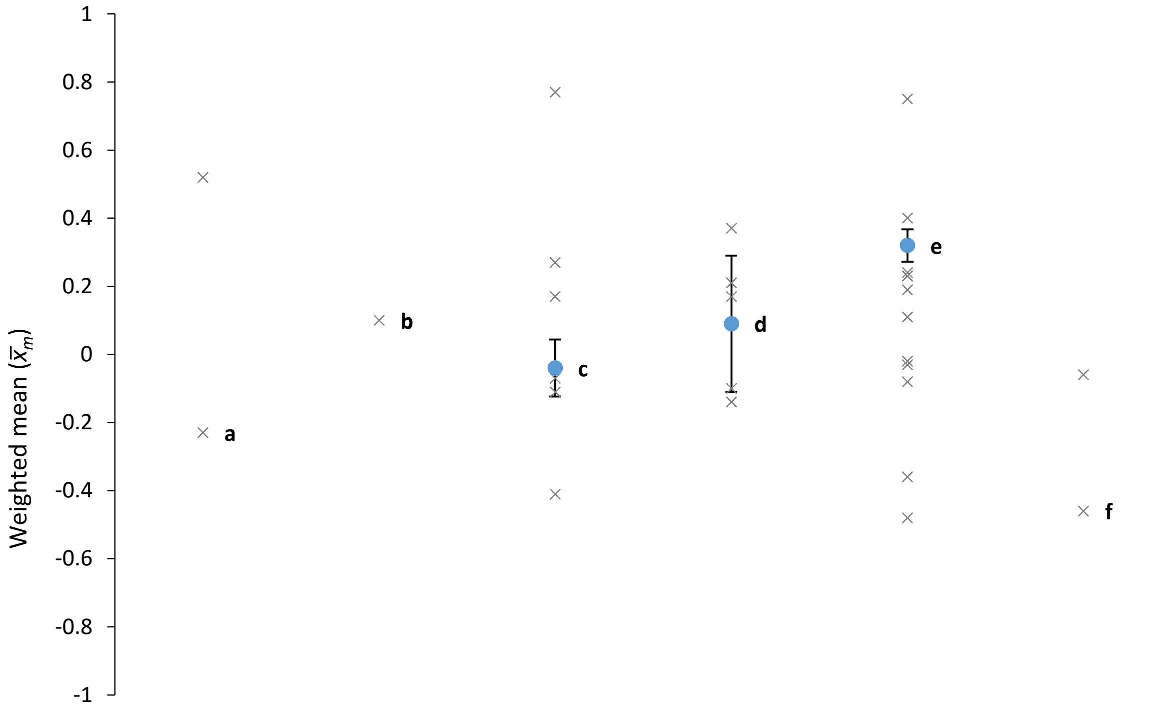


Fig 12. Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each domain of application, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each 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

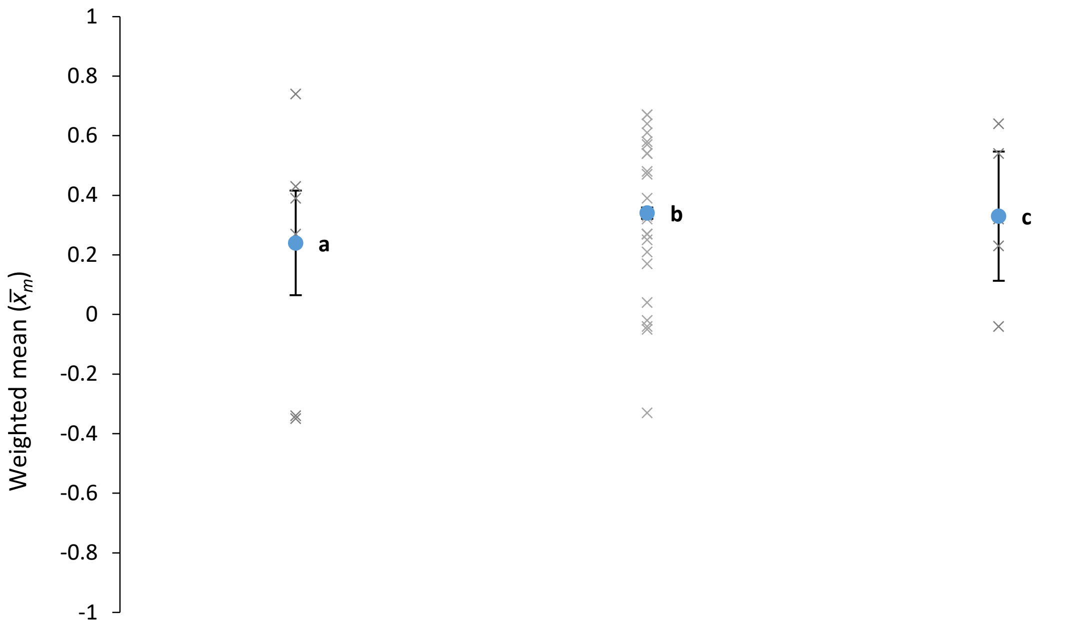


Fig 13. Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

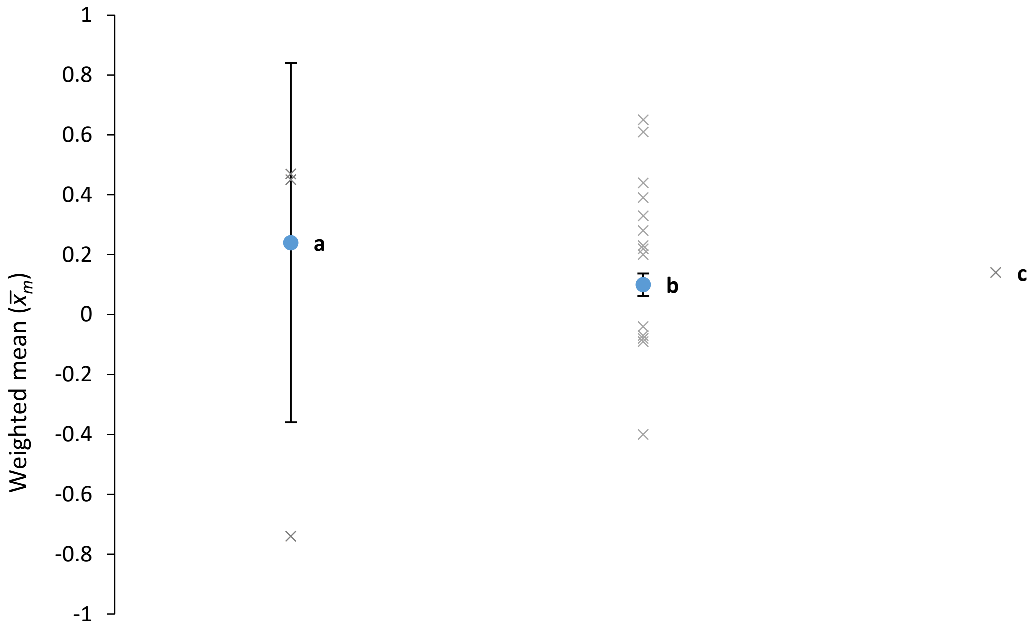


Fig 14. Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

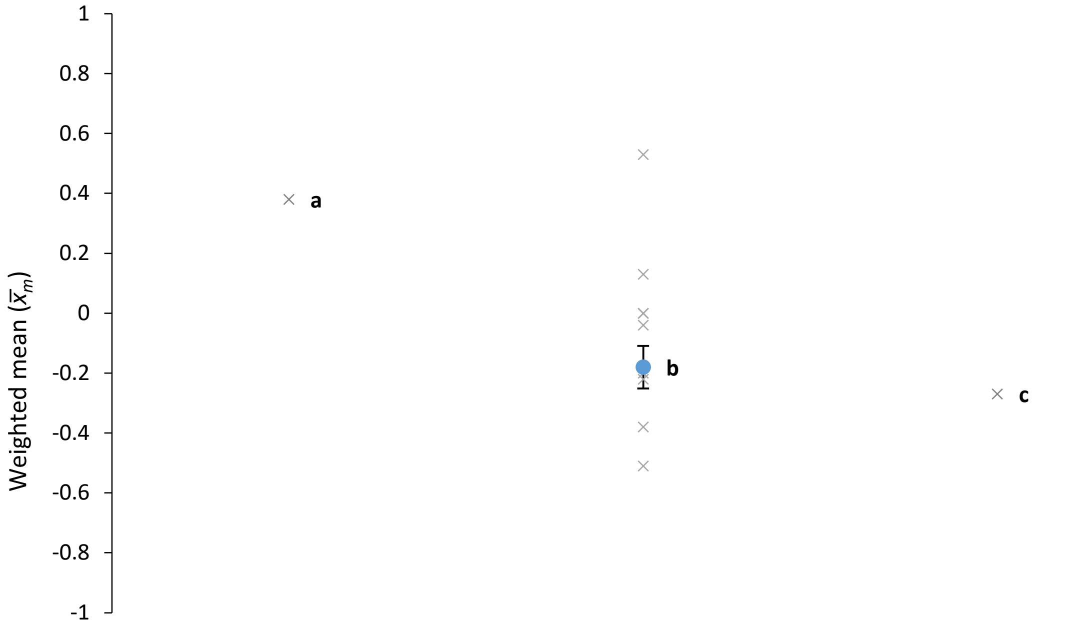


Fig 15. General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

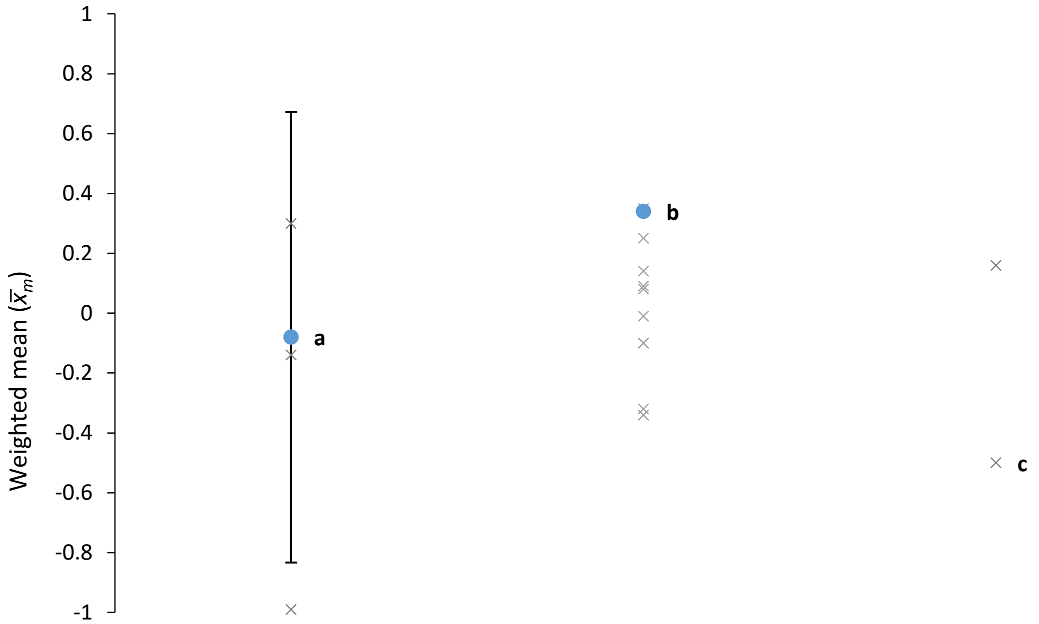


Fig 16. Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

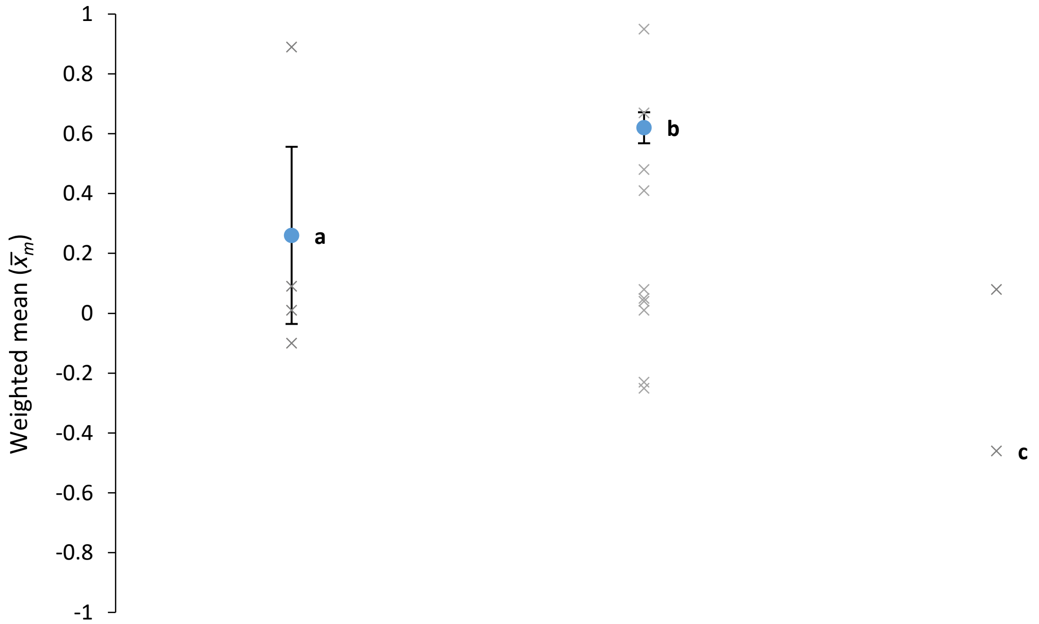


Fig 17. Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

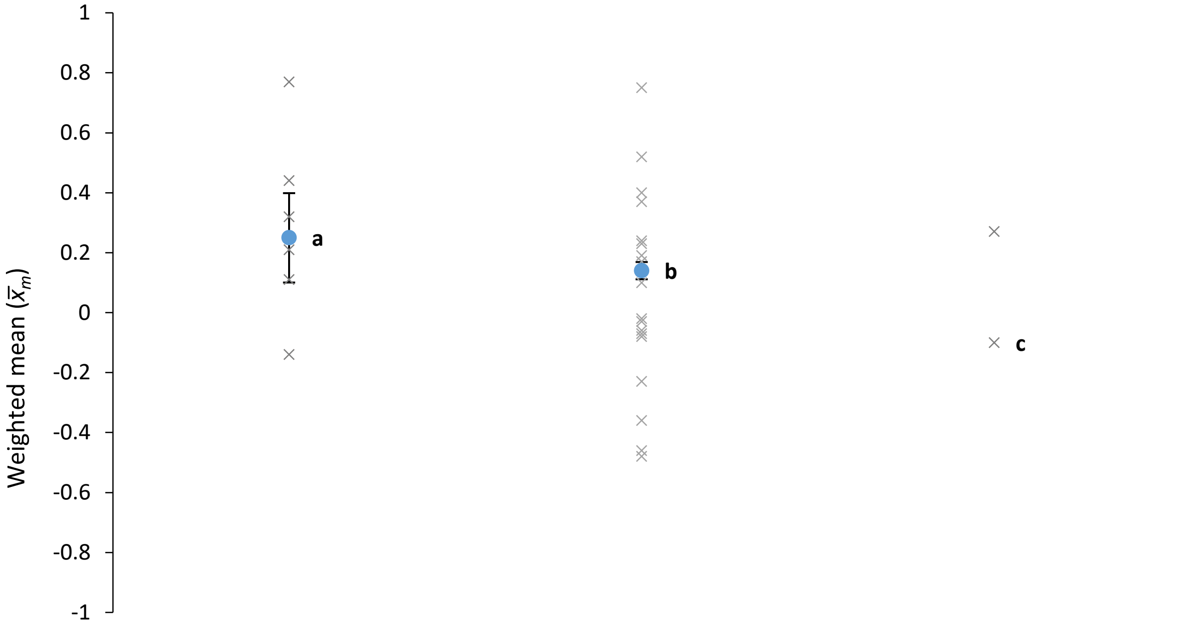


Fig 18. Trust. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) Anthropomorphic design; b) Humanoid design; and c) Non-humanoid design.

Country in which the research was conducted

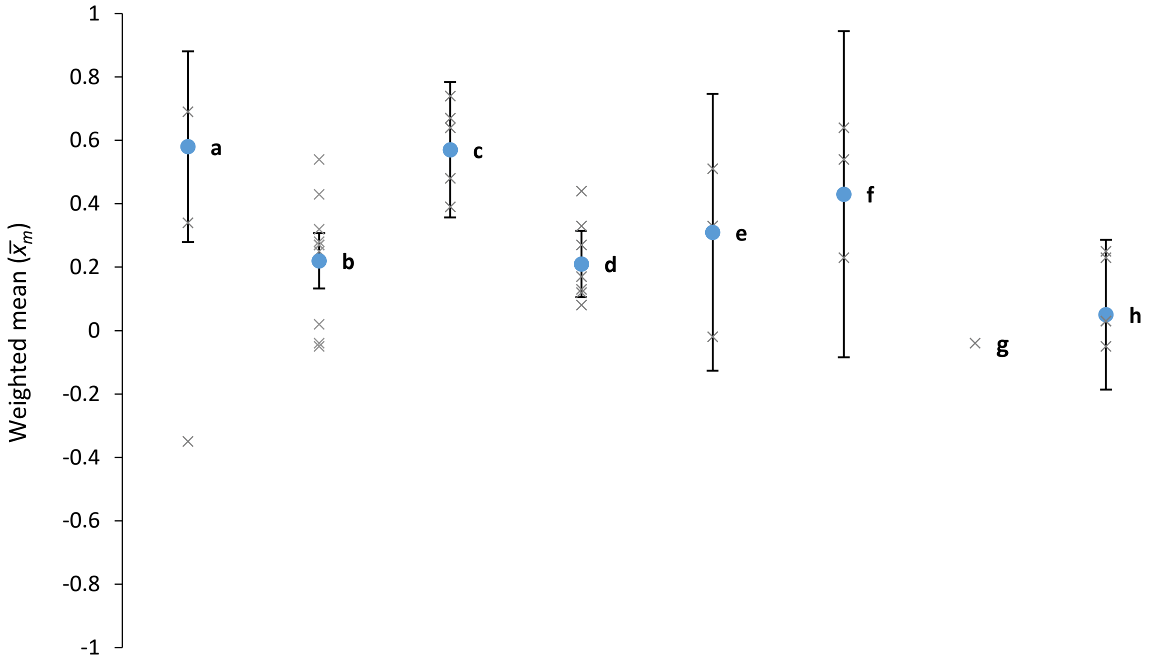


Fig 19. Affective attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; e) Netherlands; f) South Korea; g) Taiwan; and h) USA.

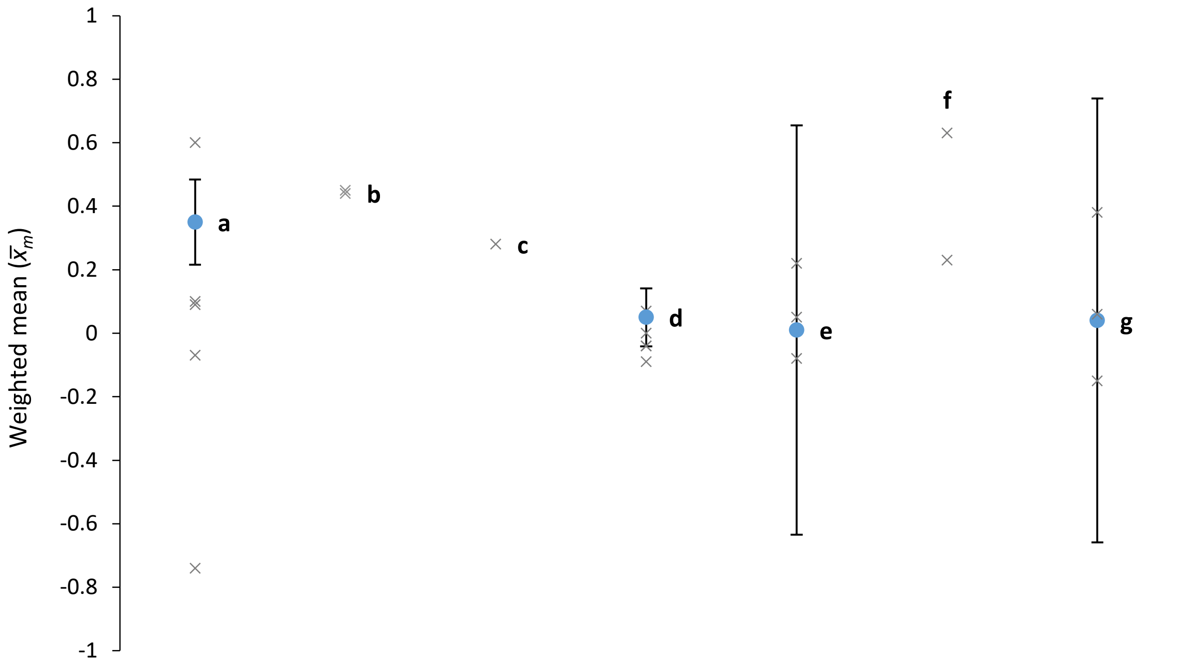


Fig 20. Cognitive attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; e) Netherlands; f) South Korea; and g) USA.

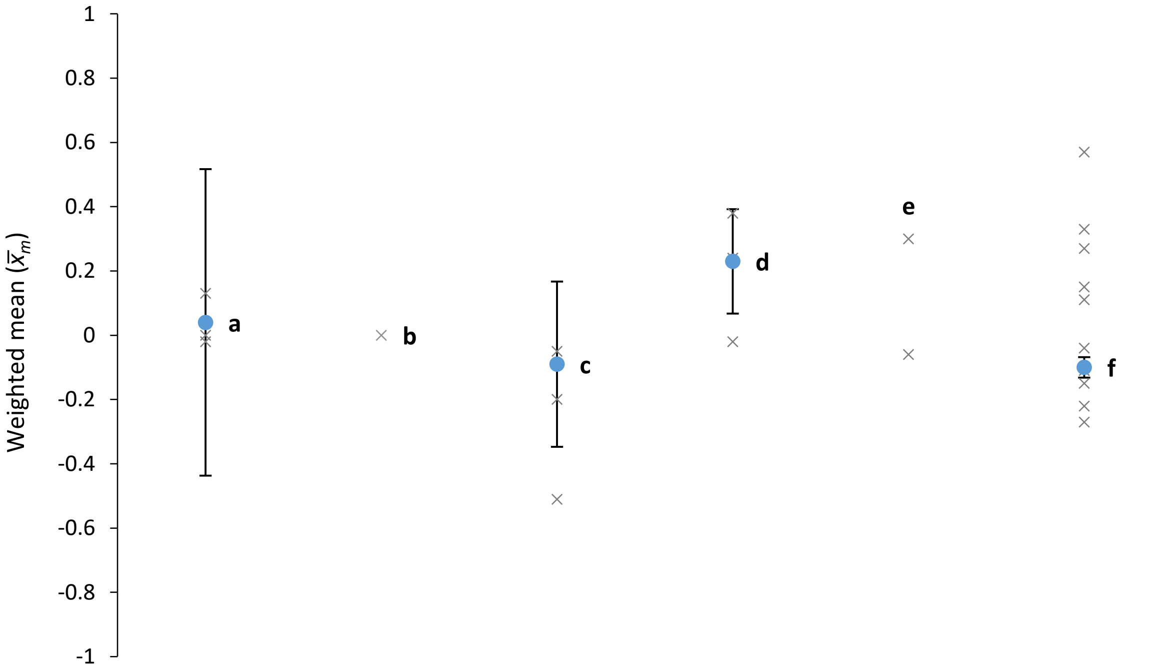


Fig 21. General attitudes. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Germany; b) Italy; c) Netherlands; d) New Zealand; e) Taiwan; and f) USA.

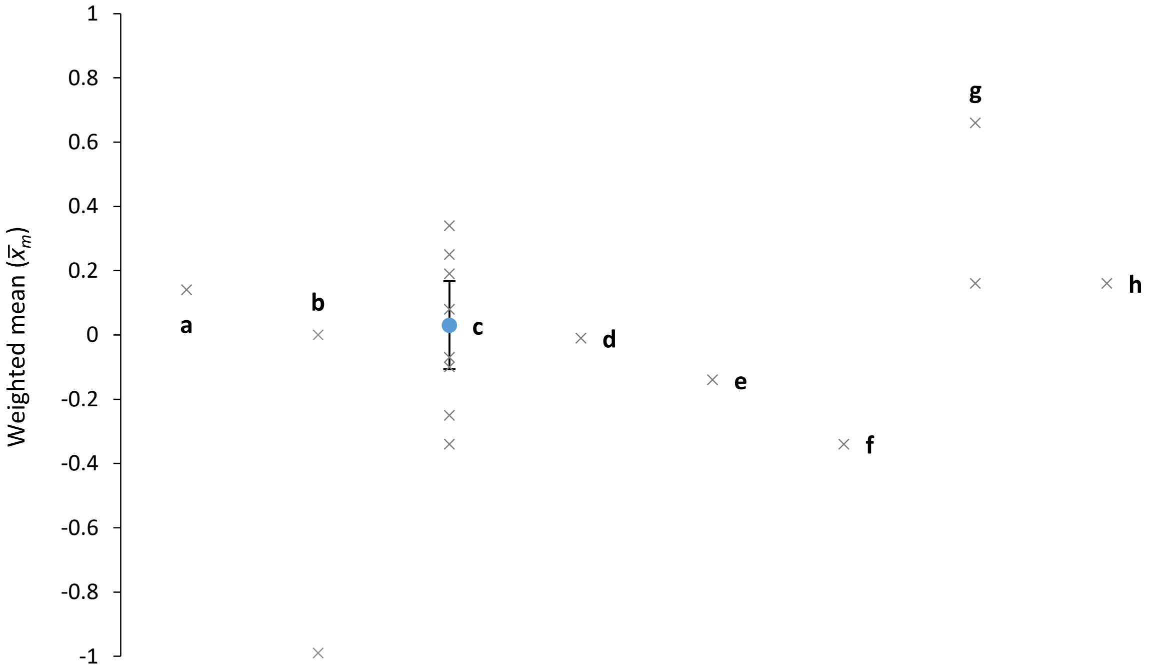


Fig 22. Acceptance. Blue data points represent the inverse-variance weighted mean ($\bar{x}_m$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x}_m}$) of the means, and the grey crosses represent the weighted means ($\bar{x}_s$) of each study in each group. a) Australia; b) France; c) Germany; d) Italy; e) Japan; f) Netherlands; g) South Korea; and h) USA.

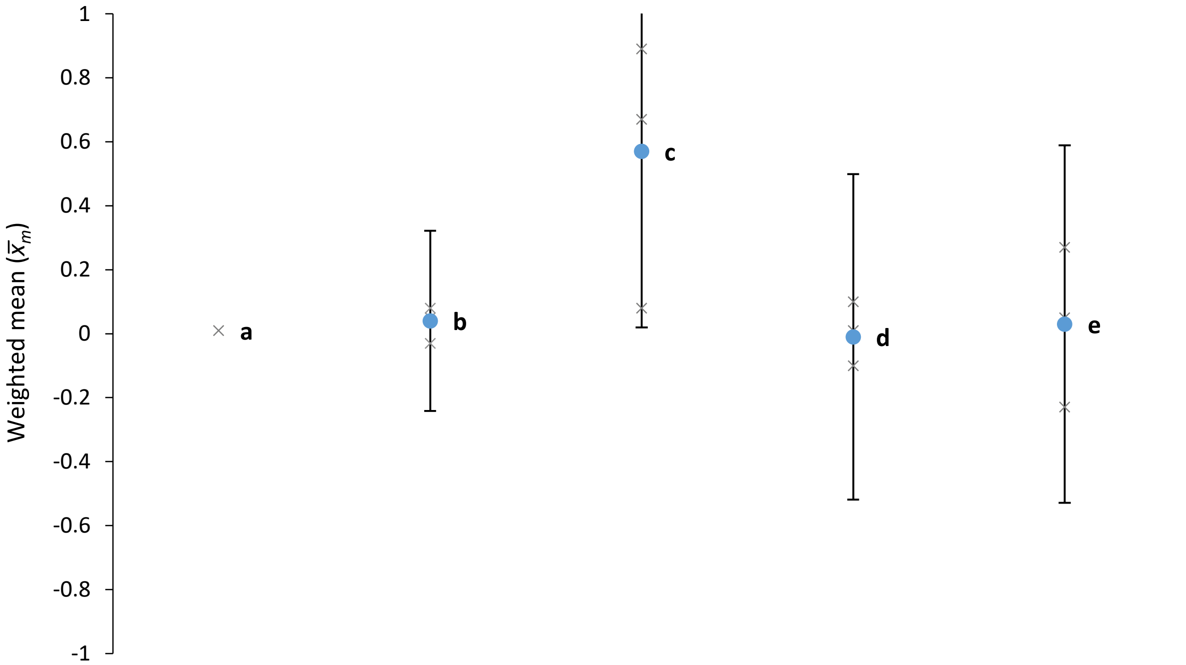


Fig 23. Anxiety. Blue data points represent the inverse-variance weighted mean ($\bar{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\bar{x_m}}$) of the means, and the grey crosses represent the weighted means ($\bar{x_s}$) of each study in each group. a) France; b) Germany; c) Italy; d) Japan; and e) Netherlands.

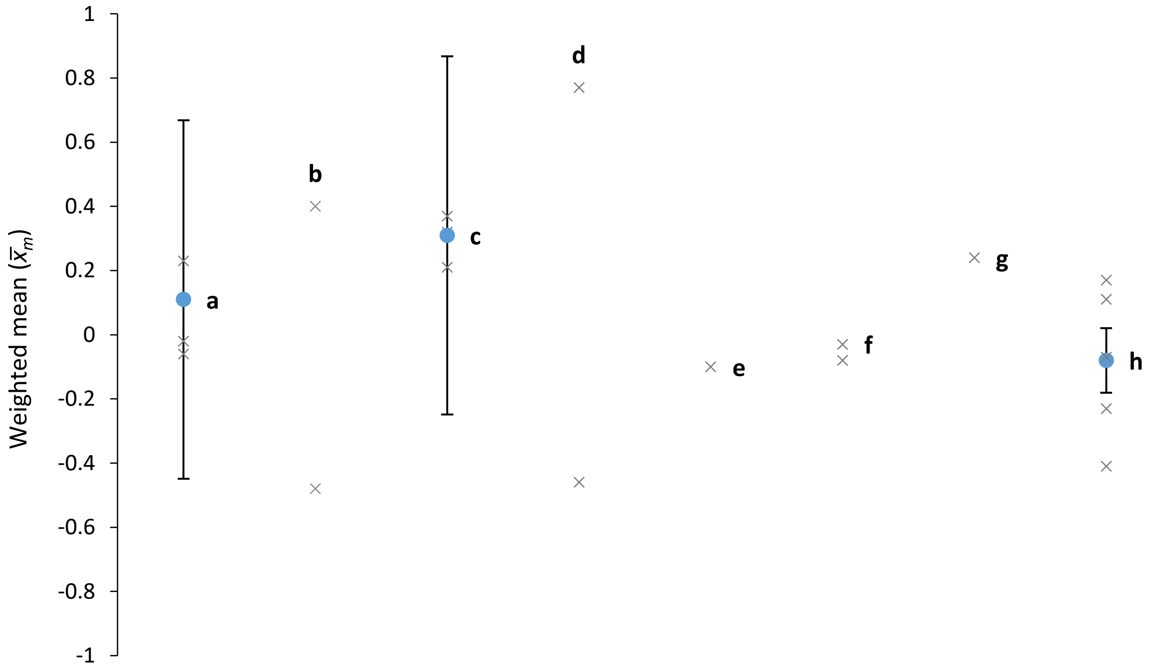


Fig 24. Trust. Blue data points represent the inverse-variance weighted mean ($\overline{x_m}$) for each type of robot design, error bars represent the weighted 95% Confidence Intervals (95% $CI_{\overline{x_m}}$) of the means, and the grey crosses represent the weighted means ($\overline{x_s}$) of each study in each group. a) Australia; b) France; c) Germany; d) Italy; e) Japan; f) Netherlands; g) South Korea; and h) USA.

Appendix D

Debrief procedure and information given to participants after each study. Studies are listed in the order in which the appear in the thesis.

Study 2

Procedure

Participants were given a physical copy of the debrief information (see below) immediately after the end of the survey. Some participants did not want a copy of the information but did receive a verbal debriefing by the researcher instead. Additionally, some participants asked questions or wished to discuss the study further which was done within the same public space at which the study took place.

Debrief Information

Research Project Title: Public knowledge of robots: A survey

Lead Researcher: Stanislava Naneva

Introduction

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have reservations due to potential job losses, loss of autonomy, privacy, and other ethical issues. While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (for example they may think robots are more advanced than what is currently possible). This can lead to negative attitudes toward robots, as well as effect interactions between people and robotic systems.

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape attitudes, acceptance, and expectations of such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Research Purpose

This study is conducted as part of the lead researcher's PhD project at the University of Sheffield.

The purpose of this exploratory survey is to investigate:

a) what the general public understands by the word ‘robot’;

b) whether fictional or non-fictional robots are more salient representatives of the ‘robot’ category;

c) whether negative (e.g., evil robots) fictional representations of robots are more salient representatives of the ‘robot’ category.

Organisation and funding

This research project has not been externally funded. This research has been organised by the lead researcher as part of their PhD at the University of Sheffield with the help of Prof. Thomas Webb and Prof. Tony Prescott.

Ethical review

This project has been ethically reviewed and approved by the Department of Psychology Ethics Committee at the University of Sheffield. More information can be found on https://www.sheffield.ac.uk/psychology/research/ethics.

Contact details

If you have any questions, queries, or suggestions regarding this research project, please contact the lead researcher, Stanislava Naneva.

Email	
Address	

If you wish to know more about robotics and the research that is currently being undertaken in Sheffield, please visit the Sheffield Robotics website: http://www.sheffieldrobotics.ac.uk/

Thank you for taking part!

Study 3

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after being asked whether they would like to be entered into a prize draw. As there was no active deception used in this study, no specific emphasis was put on any text within the main body of the information. Participants were encouraged to leave any comments at the end of the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: What do people think of when asked about robots?

Lead Researcher: Stanislava Naneva

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have reservations due to potential job losses, loss of autonomy, privacy, and other ethical issues. While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (for example they may think robots are more advanced than what is currently possible). This can lead not only to negative attitudes toward robots but also likely affects interactions between people and robotic systems. Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape attitudes, acceptance, and expectations of such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

The purpose of the research project

This study is conducted as part of the lead researcher's PhD project at the University of Sheffield. The overall purpose of the project is to understand the pattern of semantic associations people make in relation to robots, and if those associations are in turn related to people's attitudes. We are also interested in whether fictional or non-fictional robots are more salient representatives of the 'robot' category and if there is a relationship between the number of fictional / non-fictional robots people mention and their attitudes.

What if I no longer want my responses to be used?

If you wish to withdraw your data, please email the lead researcher with you unique Participant ID. There will be no negative consequences should you choose to withdraw and you will still be entered into the Amazon voucher draw.

What if I have further questions or wish to leave a comment?

If you have any questions, queries, or suggestions regarding this research project, please contact the lead researcher, Stanislava Naneva. You can do this in the box below or email them on [email] at a later date. Alternatively, you can email the project supervisor, Prof. Thomas Webb, on [email].

DOWNLOAD A COPY OF THIS INFORMATION HERE: Debrief Information - What do people think of when asked about robots.pdf

Please leave any comments you may have in the box below:

Study 4

Procedure

Participants were verbally debriefed by the research with the debrief information presented below. However, as participants were given the opportunity to have an informal discussion with the researcher and ask questions after the interview, some participants received more detailed information about the study, project, and robotics in general.

Debrief Information

As you already know, this project is looking at the way that people think about ‘robots’. More specifically, we are interested in what concepts people associate robots with and whether these associations shape people’s attitudes toward robots.

Based on previous research, we believe that people have very diverse ways of internally representing the concept of robots which in turn affects how they respond to robotics in general.

As such, we expect to see some overlapping themes between people who report having positive attitudes toward robots (or, alternatively, people who have negative attitudes toward robots).

Pilot Study 1

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after the final questionnaire in the study. As some participants received incorrect information about the videos they watched, effort was made to highlight both the fictional status and origin of each video via formatting and provision of external links. Participants were also asked to indicate whether they though they were being deceived in any way and were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the information, all participants answered the question about deception that followed the debrief information.

Debrief Information

Research Project Title: PILOT - Investigating the effect of perceived fictionality on acceptance of, and attitudes toward, robots

Lead Researcher: Stanislava Naneva [email]

Research project background

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have reservations due to potential job losses, loss of autonomy, privacy, and other ethical issues. While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (for example they may think robots are more advanced than what is currently possible). This can lead to negative attitudes toward robots, as well as affect interactions between people and robotic systems.
Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape attitudes, acceptance, and expectations of such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Aims

The aim of this pilot study is to test whether the perceived fictionality of robots can be manipulated via short descriptive texts. Both of the videos you watched were preceded by one of three possible descriptions, implying that: a) the robot is a fictional creation; b) the robot is a real-life robotic system; c) control - fictionality was not implied.

If it is possible to manipulate fictionality via text, the videos and descriptions will be used in the above-mentioned research project. Please email the lead researcher if you wish to know more.

The videos and robots

Participants received different information regarding the two videos. The true nature of the robots and videos is described below.

Video 1 – BUDDY

This video was a **promotional advertisement for a real-life companion robot** called BUDDY. Please note that this video was professionally filmed and edited. As such, BUDDY's abilities and functions **may not necessarily reflect reality**. To find out more about BUDDY and its developer, Blue Frog Robotics, please click on the following link: http://www.bluefrogrobotics.com/en/buddy/

The end of the video was not shown as it depicted BUDDY's creators talking about the robot. You can watch the video in its entirety by clicking on the following link: http://www.youtube.com/watch?v=51yGC3iytbY

Video 2 - ROBOT

This video was created by the researcher by selecting specific scenes from the movie 'Robot & Frank' (2012) and merging them together. The character of **ROBOT was played by two human actors**. Rachael Ma embodied the character in a robot costume and Peter Sarsgaard voiced ROBOT. To find out more about 'Robot & Frank' (2012), please visit this link: http://www.imdb.com/title/tt1990314/

Many of ROBOT's abilities are an **exaggeration** of the currently available technology. Whilst domestic and assistive robots are currently being developed, they are rarely commercially available and even more rarely part of someone's everyday life.

Although ROBOT is a fictional creation, it was based on a real-life humanoid robot called Asimo. You can find out more about Asimo here: http://www.asimo.honda.com/

Did you think you were being deceived OR that the information you received was incomplete?

□ Yes.

□ No.

□ Maybe / Not sure.

Do you have any additional comments or suggestions?

Thank you for taking part!

It would be greatly appreciated if you do not disclose any details of this pilot study to others.

Pilot Study 2

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after being asked whether they would like to be entered into a prize draw. As all participants were given incorrect information about the video they watched, effort was made to highlight both the non-fictional status and origin of the video via formatting and provision of external links. Participants were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: Investigating the effect of perceived fictionality on acceptance of, and attitudes toward, robots

Lead Researcher: Stanislava Naneva [email]

Research project background

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have concerns around potential job losses, loss of autonomy, privacy, and other ethical issues.

While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (e.g., they may think robots are more advanced than they currently are). This can lead to negative attitudes toward robots, as well as affect interactions between people and robotic systems.

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape peoples' beliefs about such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Aims

With the above context in mind, the aim of this study was to test whether it is possible to manipulate whether people believe that robots depicted in visual media (i.e., a video) are fictional (i.e., made up) or non-fictional (i.e., real). This is why you received one of two possible sets of information - one implying that the video depicted something fictional and the other implying that the video depicted something real, or non-fictional.

If it is possible to manipulate whether people believe that a robot depicted in a video is real or fictional, then we plan to use this procedure to investigate whether and how such beliefs influence peoples' attitudes toward the robots that they see, and advanced robotics in general. If you have any questions, then please email the lead researcher.

The video and robots

As noted above, you received one of two possible sets of information. One half of the sample was told that the video was an excerpt from a documentary (i.e., non-fictional media) meant to be shown at the next Sheffield Doc/Fest, while the other half of the sample were told that it was an excerpt from a short fiction film (fictional media) meant to be shown at the next Festival of Arts and Humanities.

In fact, neither of these statements were true. The video that you watched is actually an excerpt from the official video of a, now concluded, EU funded project called Robot-Era. You can find out more about this project here and watch the full video on YouTube.

Disclaimers

This research project is not affiliated with Sheffield Doc/Fest, the Festival of Arts and Humanities, or the Robot-Era Project.

Do you have any additional comments or suggestions?

Thank you for taking part!

It would be greatly appreciated if you do not disclose any details of this pilot study to others.

Pilot Study 3

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after being asked whether they would like to be entered into a prize draw. As all participants were given incorrect information about the video they watched, effort was made to highlight both the non-fictional status and origin of the video via formatting and provision of external links. Participants were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: Investigating the effect of perceived fictionality on acceptance of, and attitudes toward, robots

Lead Researcher: Stanislava Naneva [email]

Research project background

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have concerns around potential job losses, loss of autonomy, privacy, and other ethical issues.

While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (e.g., they may think robots are more advanced than they currently are). This can lead to negative attitudes toward robots, as well as affect interactions between people and robotic systems.

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape peoples' beliefs about such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Aims

With the above context in mind, the aim of this study was to test whether it is possible to manipulate whether people believe that robots depicted in visual media (i.e., a video) are fictional (i.e., made up) or non-fictional (i.e., real). This is why you received one of two possible sets of information - one implying that the video depicted something fictional and the other implying that the video depicted something real, or non-fictional.

If it is possible to manipulate whether people believe that a robot depicted in a video is real or fictional, then we plan to use this procedure to investigate whether and how such beliefs influence peoples' attitudes toward the robots that they see, and advanced robotics in general. If you have any questions, then please email the lead researcher.

The video and robots

As noted above, you received one of two possible sets of information. One half of the sample was told that the video was an excerpt from a documentary (i.e., non-fictional media) meant to be shown at the next UK Robotics Conference, while the other half of the sample were told that it was an excerpt from a short fiction film (fictional media) meant to be shown at a UK Sci-fi Festival.

In fact, neither of these statements were true. The video that you watched is actually an excerpt from the official video of a, now concluded, EU funded project called Robot-Era. You can find out more about this project here and watch the full video on YouTube.

Disclaimers

This research project is not affiliated with the Robot-Era Project. The UK Robotics Conference and the UK Sci-fi Festival are not real festivals and any overlap with existing events was unintentional.

Do you have any additional comments or suggestions? If yes, please leave them here.

Thank you for taking part!

It would be greatly appreciated if you do not disclose any details of this pilot study to others.

Study 5

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after the final questionnaire in the study. As all participants were given incorrect information about the video they watched, effort was made to highlight both the non-fictional status and origin of the video via formatting and provision of external links. Participants were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: Investigating the effect of perceived fictionality on attitudes toward robots

Lead Researcher: Stanislava Naneva [email]

What was this study actually about?

You were told that the aim of this research project is to investigate how advanced technology (e.g., robots) is portrayed in either fictional or non-fictional media. **However, this information was incomplete as we did not want you to know the true purpose of the study in case it influenced your responses.**

You received one of two sets of information, one stating that the video you watched was part of a short fiction film that was to be shown at the UK Sci-fi Festival OR a part of a documentary that was to be shown at the UK Robotics Conference. In fact, neither of these statements were true. The video that you watched is actually an excerpt from the official video of a, now finished, EU funded project called Robot-Era. You can find out more about this project here and watch the full video on YouTube.

The true aim of this research was to test whether it is possible to manipulate whether people believe that robots depicted in visual media (i.e., a video) are fictional (i.e., made up) or non-fictional (i.e., real). This is why you received one of two possible sets of information - one implying that the video depicted something fictional and the other implying that the video depicted something real, or non-fictional.

Research project background

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have concerns about potential job losses, loss of autonomy, privacy, and other ethical issues.

While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (e.g., they may think robots are more advanced than they currently are). This can lead to negative attitudes toward robots, as well as affect interactions between people and robotic systems.

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape peoples' beliefs about technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Why was I not told the true purpose of the research at the outset?

The true nature of this research project could not be disclosed at the beginning as it could have affected your responses.

What if I no longer want my responses to be used?

If you wish to withdraw your data, please email the lead researcher with your unique ID [inserted automatically]. There will be no negative consequences should you choose to withdraw and **you will still be entered into the Amazon voucher draw (if you so wish).**

What if I am unhappy about not being told the true purpose of the research?

You are encouraged to discuss any concerns that you may have with the lead researcher. You can do this by emailing them at [email]. Alternatively, you can email the project supervisor, Prof. Thomas Webb, on [email]. If you have any complaints or additional concerns, please email the project supervisor or the Head of the Department, [Name], on [email].

Do you have any additional comments or suggestions? If yes, please leave them here.

This is the end of the study.

Thank you for taking part!

Pilot Study 4

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after being asked whether they would like to be entered into a prize draw. Participants were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: What do people think of robots?

Lead Researcher: Stanislava Naneva

Research project background

Robots are currently being developed with the aim of assisting people in their day-to-day life. However, there are many concerns surrounding the design of such robots and the extent to which people will accept their assistance. For example, many people have concerns around potential job losses, loss of autonomy, privacy, and other ethical issues.

While some of these reservations are well founded, research suggests that some people hold unrealistic expectations about robots (e.g., they may think robots are more advanced than they currently are). This can lead to negative attitudes toward robots, as well as affect interactions between people and robotic systems.

Given that most people rarely come into contact with advanced robotics, it is likely that fictional and media representations of such robots shape peoples' beliefs about such technology. Needless to say, this can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology.

Aims

With the above context in mind, the aim of this study was to select five fictional and five non-fictional robots which will be shown to participant in a subsequent study with the aim of investigating the potential impact of fictional representations of robots on people's attitudes.

Contact details

If you have any questions, queries, or suggestions regarding this research project, then please contact the lead researcher, Stanislava Naneva.

Email [email]

Alternatively, contact the project supervisor, Prof. Thomas Webb.

Email [email]

Do you have any additional comments or suggestions? If yes, please leave them here.

Thank you for taking part!

Study 6

Procedure

Participants were presented with the debrief information (see below) online via Qualtrics immediately after being asked whether they would like to be entered into a prize draw. As participants were not made aware of the true purpose of the study, effort was made to highlight this information via formatting. Participants were encouraged to leave any comments regarding the study immediately below the debrief information. Although it was not possible to say exactly how many participants read the debrief information, nearly all participants visited the page containing the debrief information.

Debrief Information

Research Project Title: Does priming people with fictional or non-fictional robots have an effect on their attitudes toward robots?

Lead Researcher: Stanislava Naneva [email]

What was this study actually about?

You were told that the aim of this research project is to investigate people's preferences for the design of humanoid robots. **However, this information was incomplete as we did not want you to know the true purpose of the study in case it influenced your responses.**

In fact, we were not interested in your preferences for the design of robots but in your attitudes toward robots in general and if those attitudes can be changed by showing you different types of robots. In this case, we wanted to know whether people's attitudes toward robots would differ depending on whether participants were shown images of fictional robots or images of non-fictional robots.

A recent study by Thellman and Ziemke (2017) provided some evidence that people's general attitudes toward robots can be influenced by showing them different images of robots prior to measuring their attitudes. Given that most people rarely come into contact with advanced robotics, fictional and media representations of robots could play a role in shaping people's attitudes. This can be problematic as portrayals of robots in fiction rarely reflect the reality of current technology. Therefore, this particular study aims to investigate the potential impact of fiction on people's attitudes toward robots.

Why was I not told the true purpose of the research at the outset?

The true nature of this research project could not be disclosed at the beginning as it could have affected your responses.

What if I no longer want my responses to be used?

If you wish to withdraw your data, please email the lead researcher with your unique ID [inserted automatically]. There will be no negative consequences should you choose to withdraw and **you will still be entered into the Amazon voucher draw (if you so wish).**

What if I am unhappy about not being told the true purpose of the research?

You are encouraged to discuss any concerns that you may have with the lead researcher. You can do this by emailing them at [email]. Alternatively, you can email the project supervisor, Dr Thomas Webb, on [email]. If you have any complaints or additional concerns, please email the project supervisor or the Head of the Department, [Name], on [email].

Do you have any additional comments or suggestions? If yes, please leave them here.

This is the end of the study.

Thank you for taking part!

Appendix E

A list of all entities mentioned by participants in Study 2, the number (n) of times they were mentioned, and a brief description.

Entity	*n*	Description
Other	10	Any listed entity that was coded under the *Other* category.
Robots from Star Wars	9	Robots found in the fictional Star Wars movie franchise (e.g., R2D2 and C3P0).
Industrial robots	7	A class of non-fictional robots used in industry (predominantly in car manufacturing).
Robots from Robot Wars	4	A British TV show focusing on an arena-style fighting competition between custom-built robots. Although the show can be considered more entertaining than informative, the robots are non-fictional.
Asimo	3	A non-fictional bipedal humanoid robot created by the Honda company.
Marvin the Paranoid Android	3	A fictional character in 'The Hitchhiker's Guide to the Galaxy' book series by Douglas Adams and subsequent TV, movie, and radio adaptations.
Robby the Robot	3	A fictional character from the movie 'Forbidden Planet'. Robby was later adapted and appeared in multiple TV series.
Robotic hoover	3	A semi-autonomous vacuum cleaner. The most well know example of which is the Roomba model created by the iRobot company.
Drone	2	A class of non-fictional robots that can fly autonomously or, more commonly, by being remotely controlled. Used by civil and private organisations for a variety of jobs.

Entity	*n*	Description
Japanese robots	2	An ambiguous category. May refer to non-fictional robots developed by Japanese companies or to fictional robots in visual media.
K-9	2	A fictional character from the British TV series ‘Doctor Who’. This robot has a dog-like appearance.
RoboCop	2	A fictional character in multiple movies of the same name.
Robots from ‘I, Robot’	2	Fictional characters from the movie ‘I, Robot’ or, alternatively, fictional characters from Isaac Asimov's short-story collection of the same name.
WALL-E	2	A fictional character from the animated movie of the same name.
Alpha-Go	1	Not technically a robot. AlphaGo is a programme built to play the board game Go.
Baxter	1	A non-fictional industrial robot created by the Rethink Robotics company.
Cybermen	1	Fictional characters from the British TV series ‘Doctor Who’. Not technically robots but have a humanoid robot appearance.
Darth Vader	1	A fictional character from Star Wars movie franchise. Not technically a robot.
Data	1	A fictional character form the ‘Star Treck’ franchise. Technically an android, not a robot.
Deep Blue	1	Not technically a robot. A chess-playing computer developed by the IBM company.
Gort	1	A fictional humanoid character from the movie ‘The Day the Earth Stood Still’ and its remake.
Johnny 5	1	A fictional character from the movie ‘Short Circuit’.
K1 (also known as K2 & K3)	1	Fictional humanoid robots from the British TV series ‘Doctor Who’.

Entity	*n*	Description
KOALA	1	A non-fictional robot created by the K-Team company.
Maschinenmensch	1	A fictional character form the movie ‘Metropolis’.
Metal Mickey	1	A fictional character from a TV series of the same name.
Nao	1	A non-fictional humanoid robot created by the SoftBanks Robotics company
Robot from The Lost Planet	1	Could refer to one of two fictional robots characters, R-4 or R-9, from the science fiction serial, ‘The Lost Planet’.
Rosie the Robot Maid	1	A fictional character from the animated TV series, ‘The Jetsons’.
Robotic car	1	The meaning is unclear. It could refer to autonomous cars or a fictional character
Terminator	1	A fictional character from the science fiction movie franchise, ‘The Terminator’.
The Iron Giant	1	A fictional character from an animated movie of the same name.
Total	72	

Appendix F

List of all the nodes in the network constructed in Study 3 and their corresponding degree (k_i), strength (s_i), normalised strength ($s_{i_{norm}}$), and average valence ($M_{valence}$).

Node	k_i	s_i	$s_{i_{norm}}$	$M_{valence}$
advanced	10	15	0.97	0.67
algorithm	7	7	0.45	0.00
android	6	6	0.39	0.00
apocalypse	7	8	0.52	-1.00
artificial	28	41	2.65	0.00
AI	34	56	3.62	0.16
Asimov	7	8	0.52	0.00
assistance	7	7	0.45	1.00
automated	18	22	1.42	0.00
automatic	9	10	0.65	0.50
automation	11	14	0.90	0.40
autonomous	6	6	0.39	0.00
beep	4	5	0.32	1.00
book	5	6	0.39	0.50
calculating	9	9	0.58	0.33
cartoons	6	6	0.39	0.00
clean	6	6	0.39	0.50
clever	22	27	1.74	1.00
cold	17	21	1.36	-0.67
complex	7	7	0.45	1.00
complicated	7	7	0.45	0.00
computer	23	34	2.19	0.00
control	7	8	0.52	0.50
cool	5	6	0.39	1.00
cute	6	6	0.39	1.00
cyborg	5	6	0.39	-0.50

Node	k_i	s_i	$s_{i_{norm}}$	$M_{valence}$
dangerous	6	7	0.45	-1.00
development	6	6	0.39	1.00
efficient	15	18	1.16	1.00
electricity	5	6	0.39	0.00
electronics	13	15	0.97	0.00
emotionless	29	42	2.71	-1.00
engineering	12	15	0.97	0.50
expensive	7	7	0.45	-1.00
fast	13	15	0.97	1.00
fiction	6	7	0.45	0.00
friendly	5	6	0.39	1.00
funny	6	7	0.45	1.00
future	15	17	1.10	0.80
futuristic	5	5	0.32	1.00
grey	8	11	0.71	0.00
help	13	17	1.10	1.00
helpful	28	43	2.78	1.00
human	4	4	0.26	0.00
human-like	5	5	0.32	0.00
humanoid	9	11	0.71	0.25
intelligence	7	7	0.45	1.00
intelligent	19	22	1.42	0.83
jerky	13	14	0.90	-0.80
large	7	7	0.45	-0.50
logical	7	7	0.45	1.00
machine	30	48	3.10	0.00
mechanical	28	46	2.97	0.13
metal	58	128	8.26	0.05
metallic	16	18	1.16	0.00
movies	6	7	0.45	1.00
non-human	10	11	0.71	-0.25

Node	k_i	s_i	$s_{i_{norm}}$	$M_{valence}$
powerful	8	8	0.52	1.00
practical	7	7	0.45	1.00
programmed	14	17	1.10	0.00
progress	8	8	0.52	1.00
rigid	5	6	0.39	-0.50
robotic	7	7	0.45	0.00
scary	15	17	1.10	-1.00
science	16	20	1.29	0.29
science-fiction	19	27	1.74	0.25
silver	7	9	0.58	0.00
slow	3	4	0.26	-1.00
smart	14	16	1.03	0.83
space	11	13	0.84	0.00
technology	28	41	2.65	0.38
tool	5	5	0.32	0.50
toy	6	6	0.39	0.00
uncanny	10	12	0.77	-0.75
useful	26	32	2.07	1.00
voice	6	8	0.52	0.20
white	8	9	0.58	0.00
wires	9	13	0.84	-0.25
work	6	6	0.39	0.50

Appendix G

List of the nodes in the network constructed in Study 3 and the intramodular strength (sm_i) and normalised intramodular strength ($sm_{i_{norm}}$) of the nodes.

Node name	sm_i	$sm_{i_{norm}}$
	Module 1	
Apocalypse	4	0.40
Algorithm	6	0.60
Android	6	0.60
AI	29	2.90
Asimov	4	0.40
Beep	4	0.40
Book	4	0.40
Computer	23	2.30
Cyborg	6	0.60
Electricity	3	0.30
Fiction	6	0.60
Funny	5	0.50
Grey	6	0.60
Help	12	1.20
Human	4	0.40
Metal	46	4.60
Movies	3	0.30
Sci-fi	14	1.40
Silver	6	0.60
Space	8	0.80
Wires	11	1.10
	Module 2	
Advanced	8	0.94
Assistance	5	0.58
Automated	11	1.29
Automatic	6	0.70

Node name	sm_i	$sm_{i_{norm}}$
Clever	12	1.40
Complex	3	0.35
Cool	5	0.58
Cute	5	0.58
Development	3	0.35
Efficient	10	1.17
Friendly	5	0.58
Future	12	1.40
Futuristic	4	0.47
Helpful	22	2.57
Human-like	3	0.35
Intelligent	9	1.05
Non-human	5	0.58
Robotic	4	0.47
Scary	12	1.40
Science	8	0.94
Technology	16	1.87
Useful	20	2.34
Module 3		
Autonomous	2	0.35
Calculating	5	0.87
Cartoons	2	0.35
Clean	4	0.69
Cold	10	1.73
Complicated	3	0.52
Emotionless	18	3.11
Expensive	4	0.69
Logical	2	0.35
Mechanical	14	2.42
Practical	4	0.69
Programmed	7	1.21

Node name	sm_i	$sm_{i_{norm}}$
Rigid	4	0.69
Slow	4	0.69
Smart	10	1.73
Voice	4	0.69
White	4	0.69
Work	3	0.52
Module 4		
Artificial	12	2.12
Automation	5	0.88
Control	6	1.06
Electronics	7	1.23
Engineering	9	1.59
Humanoid	7	1.23
Intelligence	3	0.53
Jerky	3	0.53
Metallic	7	1.23
Progress	5	0.88
Tool	2	0.35
Toy	2	0.35
Module 5		
Dangerous	6	1.20
Fast	5	1.00
Large	4	0.80
Machine	8	1.60
Powerful	4	0.80
Uncanny	3	0.60

Appendix H

A visual breakdown of each module in the network constructed in Study 3.

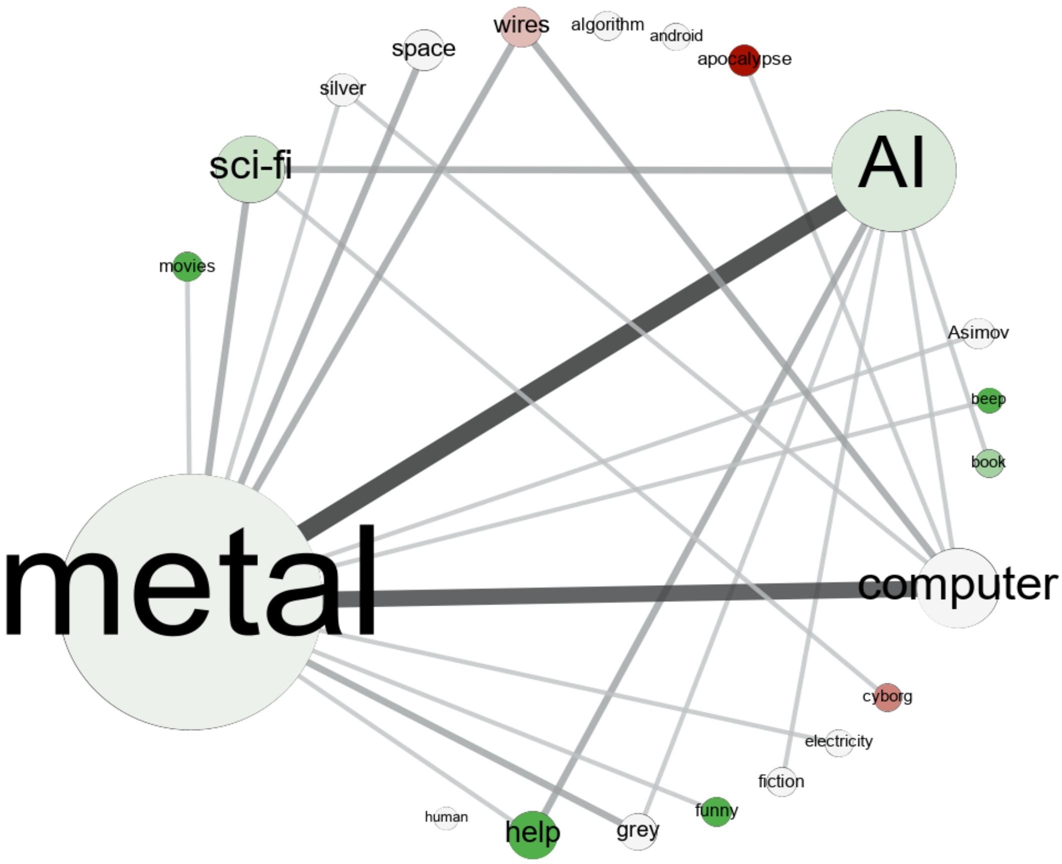


Figure E1. Visualisation of all the nodes/associations in Module 1 in alphabetical order. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

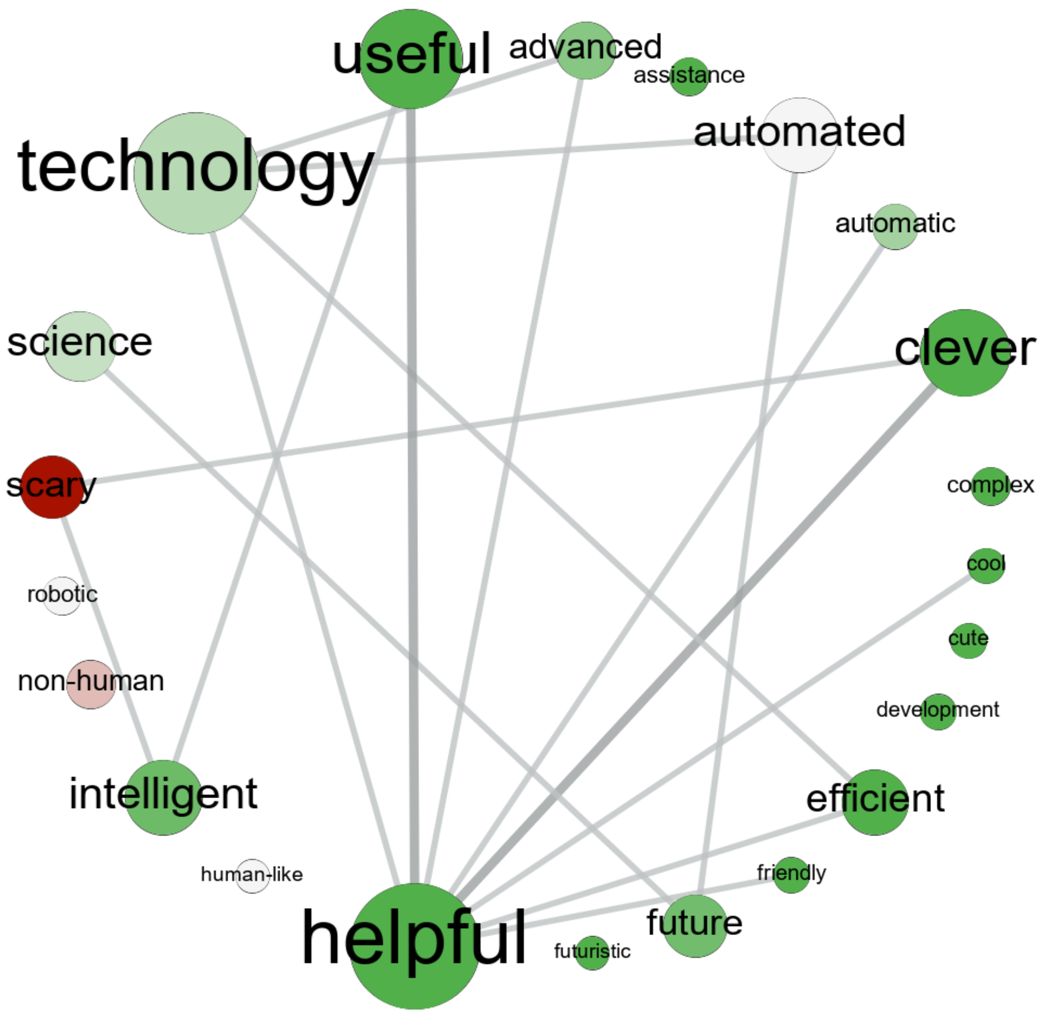


Figure E2. Visualisation of all the nodes/associations in Module 2 in alphabetical order. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

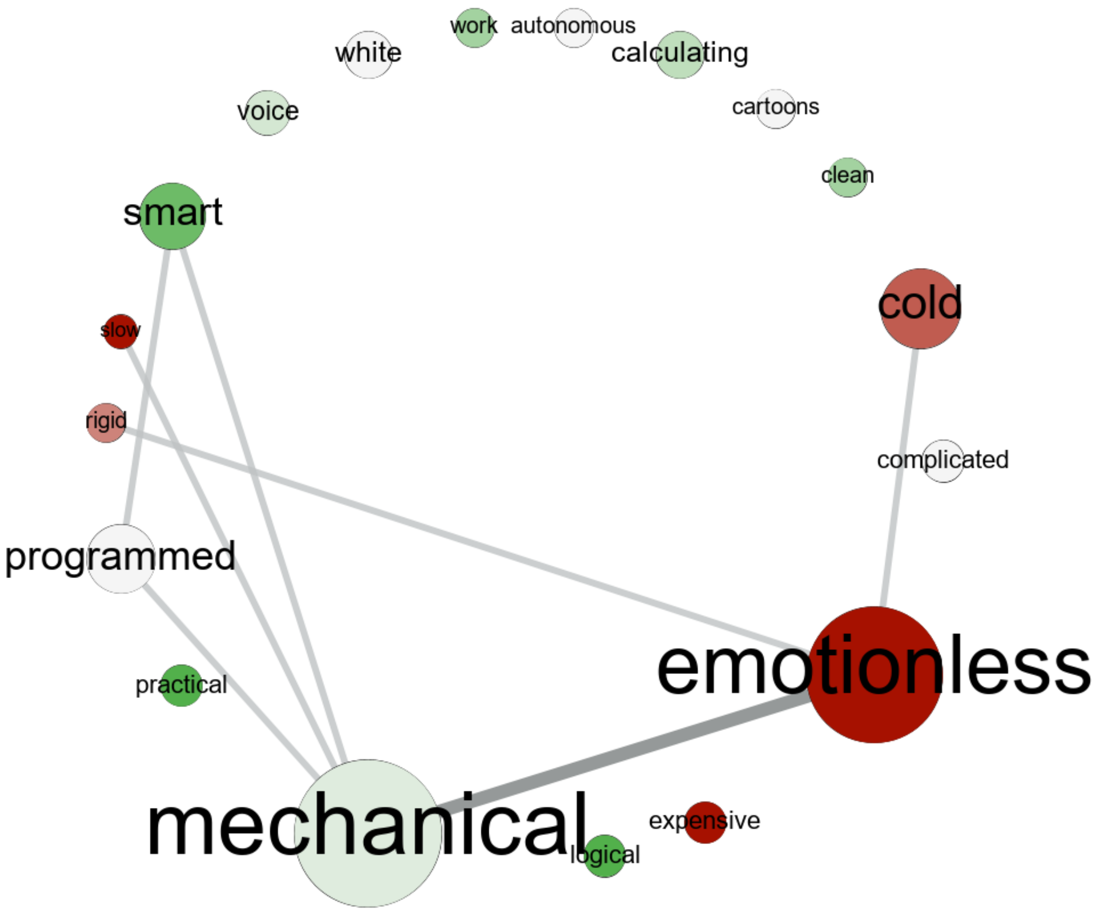


Figure E3. Visualisation of all the nodes/associations in Module 3 in alphabetical order. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

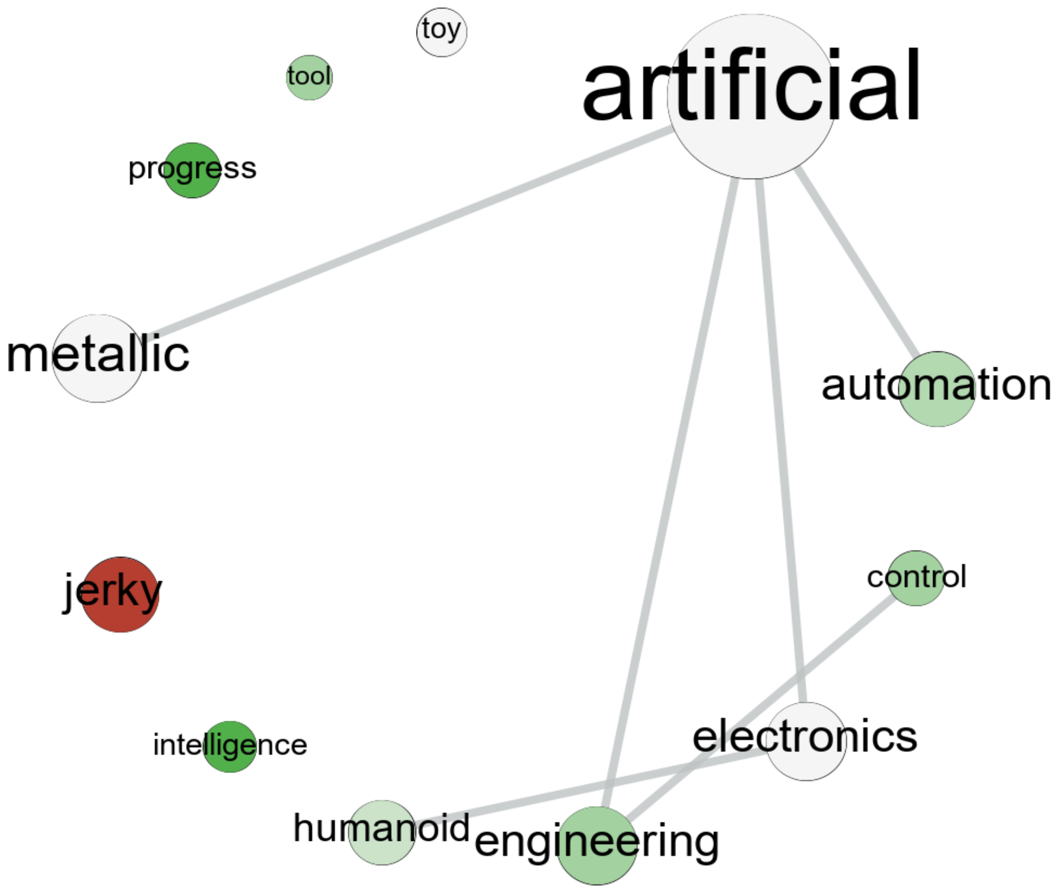


Figure E4. Visualisation of all the nodes/associations in Module 4 in alphabetical order. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

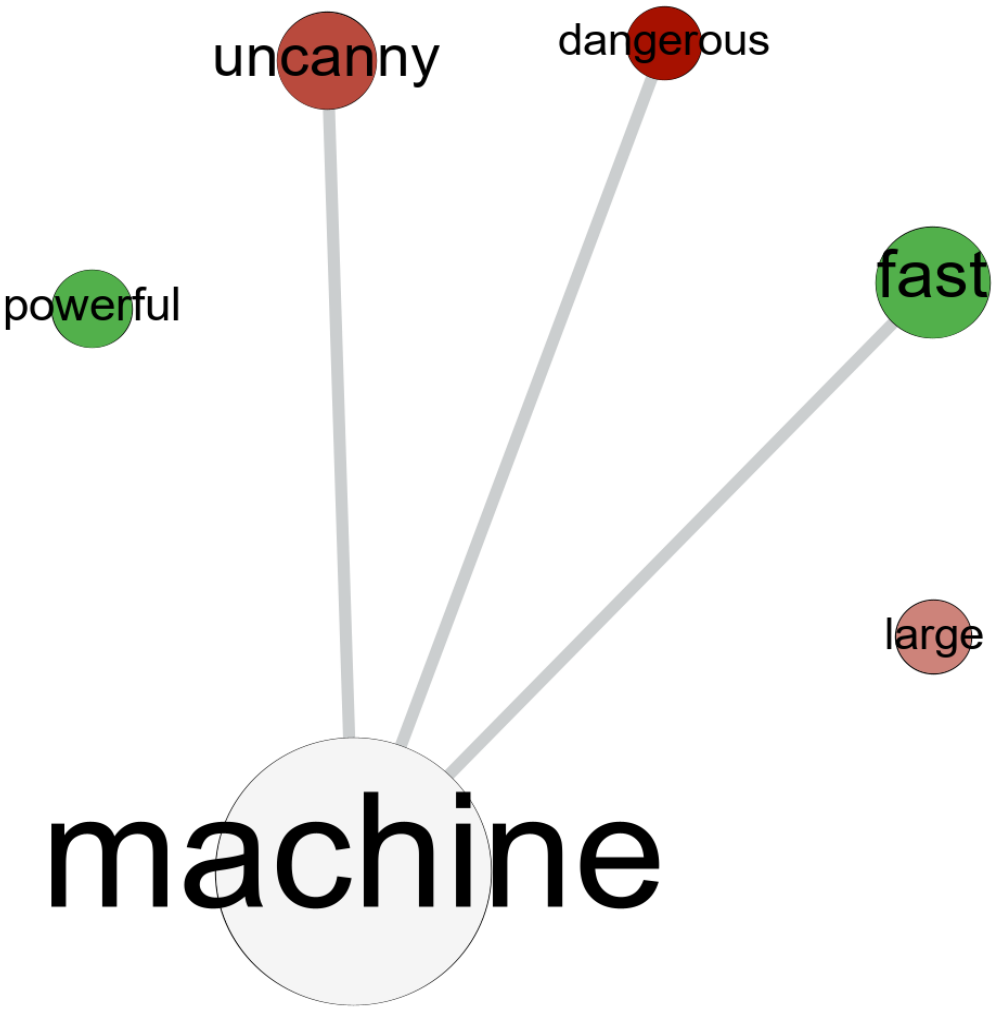


Figure E5. Visualisation of all the nodes/associations in Module 5 in alphabetical order. Size of the nodes represents the node strength and the size of the edges is determined by their weight. Edges with a weight of less than 3 have been removed for better visualisation. Colour gradient indicates the average valence of each association such that the darkest green indicates the most positive association (+1) and the darkest red indicates the most negative associations (-1).

Appendix I

A list of all entities mentioned by participants in the Study 3 survey detailing the number of mentions (n) and the percentage of the data which the mentions cover. The asterisk (*) represents entities that were merged together into a single category (typically the source material) as it was unclear to which robot characters participants were referring to.

Robot	*n*	Data coverage (%)	Origin	Fictional status
R2-D2	37	8.75%	‘Star Wars' (1977 - 2019) film franchise	Fictional
Wall-E	31	7.33%	‘Wall-E' (2008) animated film	Fictional
C-3P0	24	5.67%	‘Star Wars' (1977 - 2019) film franchise	Fictional
*Terminator	20	4.73%	‘Terminator' (1984 - 2019) film franchise	Fictional
Unidentified fictional robots	17	4.02%	Various	Fictional
*Transformers	14	3.31%	‘Transformers' (2007 - 2018) film franchise	Fictional
Unidentified non-fictional robots	15	3.55%	Various	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Johnny 5	11	2.60%	‘Short Circuit' (1986) film	Fictional
Sophia	11	2.60%	A humanoid robot developed by Hanson Robotics (2016)	Non-fictional
Manufacturing robots	10	2.36%	Various	Non-fictional
*RoboCop	9	2.13%	‘RoboCop' (1987 - 2014) film franchise	Fictional
Robotic hoover	9	2.13%	Various	Non-fictional
Bender	8	1.89%	‘Futurama' (1999 - 2013) animated TV series	Fictional
Dalek	8	1.89%	‘Doctor Who' (1963 - 2019) TV series	Fictional
*I, Robot	8	1.89%	‘I, Robot' (2014) film & Isaac Asimov's short-story collection (1950) of the same name	Fictional
K9	8	1.89%	‘Doctor Who' (1963 - 2019) TV series	Fictional
Alexa	7	1.65%	A virtual assistant developed by Amazon (2014)	Non-fictional
HAL 9000	7	1.65%	Arthur C. Clarke's Space Odyssey (1968 - 1997) book series & films of the same name	Fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
EVE	6	1.42%	‘Wall-E' (2008) animated film	Fictional
*Robots	6	1.42%	‘Robots' (2005) film	Fictional
Baymax	5	1.18%	‘Big Hero 6' (2014) animated film	Fictional
Data	5	1.18%	‘Star Trek' TV series (1966 - 2019) and film franchise (1979 - 2016)	Fictional
Deep Blue	5	1.18%	A chess-playing computer developed by IBM (1995)	Non-fictional
NASA rovers	5	1.18%	Various, developed by The National Aeronautics and Space Administration	Non-fictional
BB-8	4	0.95%	‘Star Wars' (1977 - 2019) film franchise	Fictional
Robby the Robot	4	0.95%	‘Forbidden Planet' (1956) film	Fictional
Robot Wars	4	0.95%	‘Robot Wars' (1998 - 2018) TV series and robotic competition	Non-fictional
Atlas	3	0.71%	A bipedal humanoid robot developed by Boston Robotics (2013)	Non-fictional
Ava	3	0.71%	‘Ex Machina' (2014) film	Fictional
Computer	3	0.71%	Various	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Humans	3	0.71%	‘Humans' (2015 - 2018) TV series	Fictional
Robots seen at a conference	3	0.71%	Various	Non-fictional
Siri	3	0.71%	A virtual assistant developed by Apple (2011)	Non-fictional
Unidentified other robots	3	0.71%	Various	Other
Asimo	2	0.47%	A bipedal humanoid robot developed by Honda (2000)	Non-fictional
CHAPPiE	2	0.47%	‘Chappie' (2015) film	Fictional
Children's toys	2	0.47%	Various	Non-fictional
Google	2	0.47%	-	Non-fictional
KITT	2	0.47%	‘Knight Rider' (1982 - 2008) TV series	Fictional
Mac and C.H.E.E.S.E	2	0.47%	‘Friends' (1994 - 2004) TV series	Fictional
Marvin The Paranoid Android	2	0.47%	Douglas Adams' 'The Hitchhiker's Guide to the Galaxy' book series and film	Fictional
Medical robots	2	0.47%	Various	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
MiRo	2	0.47%	A zoomorphic robot developed by Consequential Robotics	Non-fictional
Overwatch (general)	2	0.47%	‘Overwatch' (2016) video game developed and published by Blizzard Entertainment	Fictional
Pepper	2	0.47%	A humanoid robot developed by SoftBank Robotics (2014)	Non-fictional
Robots (Aliens)	2	0.47%	‘Aliens' (1986) film	Fictional
Robots seen at TUoS	2	0.47%	Various	Non-fictional
SpotMini	2	0.47%	A zoomorphic robot developed by Boston Dynamics (2016)	Non-fictional
Zoomorphic robot (health care)	2	0.47%	Various	Non-fictional
2-XL	1	0.24%	An educational toy robot by Mego Corporation (1978 - 1981)	Non-fictional
Ai-Da	1	0.24%	A humanoid robot developed by Engineered Arts (2019)	Non-fictional
Autonomous cars	1	0.24%	Various	Non-fictional
Bicentennial Man	1	0.24%	‘Bicentennial Man' (1999) film and Isaac Asimov's novella (1976)by the same name	Fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Black Mirror (general)	1	0.24%	‘Black Mirror' (2011 - 2019) TV series	Fictional
Bleep and Booster	1	0.24%	‘Bleep and Booster' (1964 - 1977) animated TV series	Fictional
Blender	1	0.24%	Various	Non-fictional
Bomb disposal robot	1	0.24%	Various	Non-fictional
Bomb disposal robot (Brooklyn 99)	1	0.24%	‘Brooklyn Nine-nine' (2013 - 2019) TV series	Fictional
Car parking ticket machine	1	0.24%	Various	Non-fictional
CHiP	1	0.24%	A zoomorphic robot developed by WowWee	Non-fictional
Chuck (Mork & Mindy)	1	0.24%	‘Mork & Mindy' (1978 - 1982) TV series	Fictional
Cybermen	1	0.24%	‘Doctor Who' (1963 - 2019) TV series	Fictional
Darth Vader	1	0.24%	‘Star Wars' (1977 - 2019) film franchise	Fictional
DeepMind	1	0.24%	A group of programs developed by Google	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Delivery robots	1	0.24%	Various	Non-fictional
Dingbot	1	0.24%	A toy robot developed by Tomy	Non-fictional
Dishwasher	1	0.24%	Various	Non-fictional
Disnely Land robot	1	0.24%	Various	Non-fictional
Donald Trump	1	0.24%	-	Other
Ecci	1	0.24%	A humanoid robot developed by the University of Zurich	Non-fictional
Echo	1	0.24%	A smart speaker developed by Amazon (2014)	Non-fictional
Eric	1	0.24%	A humanoid robot developed by William Richards and Alan Reffell (1928)	Non-fictional
Evas	1	0.24%	‘Neon Genesis Evangelion' (1995 - 2012) animated TV series and franchise	Fictional
Ferbots	1	0.24%	‘Phineas and Ferb' (2007 - 2015) animated TV series	Fictional
GERTY	1	0.24%	‘Moon' (2009) film	Fictional
Glados	1	0.24%	‘Portal' (2007) and 'Portal 2' (2011) video games	Fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Gort	1	0.24%	The Day the Earth Stood Still' (1951, 2008) film	Fictional
Gundam (general)	1	0.24%	‘Gundam' (1979 - 2019) animated franchise	Fictional
Hank (Final Space)	1	0.24%	‘Final Space' (2018 - 2019) animated TV series	Fictional
Health care robot	1	0.24%	Various	Non-fictional
HK-47	1	0.24%	‘Star Wars' (2003 - 2011) video game franchise	Fictional
Watson	1	0.24%	An AI developed by IBM (2010)	Non-fictional
iPhone	1	0.24%	A smart phone series developed by Apple	Non-fictional
Iron Giant	1	0.24%	‘The Iron Giant' (1999) animated film	Fictional
Irona	1	0.24%	‘Richie Rich' (1980 - 1984) animated TV series	Fictional
Jaegers	1	0.24%	‘Pacific Rim' (2013) film	Fictional
K.A.R.E.N	1	0.24%	‘SpongeBob SquarePants' (1999 - 2019) animated TV series	Fictional
Karakuri	1	0.24%	Various	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Kryten	1	0.24%	‘Red Dwarf' (1988 - 1999) TV series	Fictional
KUKA	1	0.24%	Industrial robotic arms developed by KUKA	Non-fictional
Loader Bot	1	0.24%	‘Borderlands' (2009 - 2019) game franchise	Fictional
Lore	1	0.24%	‘Star Trek: The Next Generation' (1987 - 1994) TV series	Fictional
Maria / Maschinenmensch	1	0.24%	‘Metropolis' (1927) film	Fictional
Mecha (Macross)	1	0.24%	‘Macross' (1982 - 2018) animated franchise	Fictional
Mega Man	1	0.24%	‘Mega Man' (1987 - 2018) video game franchise	Fictional
Metal Mickey	1	0.24%	‘The Saturday Banana' (1987) TV show	Fictional
NAO	1	0.24%	A humanoid robot developed by the SoftBank robotics (2004)	Non-fictional
Omnibot	1	0.24%	A toy robot developed by Tomy (1980s)	Non-fictional
PIP	1	0.24%	A non-humanoid robot developed by Swallow Systems	Non-fictional
R.O.B	1	0.24%	A toy robot developed by Nintendo (1985)	Non-fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Relativity Space's robot	1	0.24%	A system developed by Relativity Space	Non-fictional
Replicants	1	0.24%	‘Blade Runner' (1982) film and 'Blade Runner' (2017) film	Fictional
M3-B9 G.U.N.T.E.R.	1	0.24%	‘Lost in Space' (1965 - 1968) TV series	Fictional
Robot (Toy Story)	1	0.24%	‘Toy Story' (1999 - 2018) animated film franchise	Fictional
Robotic hand	1	0.24%	Various	Non-fictional
Robotic lawnmower	1	0.24%	Various	Non-fictional
Robots (UKS)	1	0.24%	‘Unbreakable Kimmy Schmidt' (2015 - 2019) TV series	Fictional
Robots seen at a robotics contest	1	0.24%	Various	Non-fictional
Rosie	1	0.24%	‘The Jetsons' (1962 - 1987) animated TV series	Fictional
RUR	1	0.24%	‘R.U.R.' (1920) play	Fictional
Samantha	1	0.24%	‘Her' (2013) film	Fictional

Robot	*n*	Data coverage (%)	Origin	Fictional status
Sci-fi memorabilia	1	0.24%	Various	Fictional
Self-checkout machines	1	0.24%	Various	Non-fictional
Sico	1	0.24%	‘Rocky IV' (1985) film	Fictional
SlugBot	1	0.24%	A non-humanoid robot developed by Ian Kelly	Non-fictional
Table tennis robot	1	0.24%	Various	Non-fictional
The Android (Dark Matter)	1	0.24%	‘Dark Matter' (2015 - 2017) TV series	Fictional
T-HR3	1	0.24%	A humanoid robot developed by Toyota	Non-fictional
Tipster	1	0.24%	A toy developed by WowWee (2014)	Non-fictional
Tommy the Robot	1	0.24%	A robot developed by Zyrobotics	Non-fictional
Traffic lights	1	0.24%	Various	Non-fictional
Unidentified robot (Simpsons)	1	0.24%	‘The Simpsons' (1989 - 2019) animated TV series	Fictional
Washing machine	1	0.24%	Various	Non-fictional

Appendix J

Assumption checks relevant to the analyses presented in the Results section of Pilot Study 1.

Video 1: 'BUDDY: Your Family's Companion Robot'

Realism of the robot

Shapiro-Wilk test of normality was significant for the non-fiction condition (see Table I1). There was no substantial skewness or kurtosis (indicated by *z* scores larger than ± 1.96) for any of the conditions. Further examination of Q-Q plots indicated that the scores within groups were sufficiently normally distributed. A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated.

Table I1

Normal distribution statistics for each of the three conditions (robot realism subscale) in Pilot Study 1

Condition	*W*	*df*	*p*	Z_s	Z_k
Fiction	.95	12	.587	0.42	-0.41
Non-fiction	.76	13	.002	-1.27	-0.41
Control	.93	12	.373	-1.42	1.02

Note. *W* obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Realism of the video

Shapiro-Wilk test of normality was significant for the fiction condition (see Table I2). There was no substantial skewness or kurtosis (indicated by *z* scores larger than ± 1.96) for any of the conditions. Further examination of Q-Q plots indicated that the scores within groups were sufficiently normally distributed. A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated.

Table I2

Normal distribution statistics for each of the three conditions (video realism subscale) in Pilot Study 1

Condition	*W*	*df*	*p*	Z_s	Z_k
Fiction	.81	12	.011	0.82	-1.02
Non-fiction	.86	13	.034	1.25	-0.49
Control	.89	12	.137	-1.99	1.84

Note. *W* obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Quality of the video

Shapiro-Wilk test of normality was non-significant for all conditions (see Table I3). There was no substantial skewness or kurtosis (indicated by *z* scores larger than ± 1.96) for any of the conditions. Further examination of Q-Q plots indicated that the scores within groups were sufficiently normally distributed. A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated.

Table I3

Normal distribution statistics for each of the three conditions (video quality subscale) in Pilot Study 1

Condition	*W*	*df*	*p*	Z_s	Z_k
Fiction	.89	12	.106	-0.04	-1.21
Non-fiction	.82	13	.012	-1.64	-0.22
Control	.84	12	.029	0.84	-1.11

Note. *W* obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Video 2: 'Robot & Frank'

Realism of the robot

Shapiro-Wilk test of normality was non-significant for all conditions (see Table I4). There was no substantial skewness or kurtosis (indicated by z scores larger than ± 1.96) for any of the conditions. Further examination of Q-Q plots indicated that the scores within groups were sufficiently normally distributed. A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated.

Table I4

Normal distribution statistics for each of the three conditions (robot realism subscale) in Pilot Study 1

Condition	W	df	p	Z_s	Z_k
Fiction	.87	12	.070	1.00	-0.59
Non-fiction	.86	12	.047	1.02	-0.93
Control	.91	12	.194	-0.89	-0.54

Note. W obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Realism of the video

Shapiro-Wilk test of normality was significant for the fiction condition (see Table I5). There was no substantial skewness or kurtosis (indicated by z scores larger than ± 1.96) for any of the conditions. Further examination of Q-Q plots indicated that the scores within groups were sufficiently normally distributed. A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated.

Table I5

Normal distribution statistics for each of the three conditions (video realism subscale) in Pilot Study 1

Condition	*W*	*df*	*p*	Z_s	Z_k
Fiction	.82	12	.015	-0.41	-0.56
Non-fiction	.87	12	.056	-1.12	0.43
Control	.88	12	.080	0.20	-0.80

Note. *W* obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Quality of the video

Shapiro-Wilk test of normality was significant for the non-fiction condition (see Table I6). Furthermore, the non-fiction condition was substantially leptokurtic (indicated by *z* scores larger than ± 1.96) and somewhat skewed. This is probably due to the scale items as the video was consistently rated as high quality (median = 4). A non-significant Levene's test indicated that the assumption of homogeneity of variance was not violated. Since the group sizes are roughly equal, the ANOVA's robustness should be able to cover for the moderate departure from normality.

Table I6

Normal distribution statistics for each of the three conditions (video quality subscale) in Pilot Study 1

Condition	*W*	*df*	*p*	Z_s	Z_k
Fiction	.90	12	.148	-1.70	0.88
Non-fiction	.81	12	.013	-1.83	2.23
Control	.90	12	.181	-0.61	0.89

Note. *W* obtained from Shapiro-Wilk test of normality, Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis. Skewness and kurtosis values greater than ±1.96 are considered substantially non-normal.

Appendix K

Assumption checks and treatment of the data relevant to the analyses presented in the Results section of Pilot Study 2.

Plausibility, typicality, factuality, perceptual quality, and narrative consistency scores were obtained by calculating the mean of all items in each subscale and analysed using SPSS Statistics 23. Each item (for each of the two conditions) were transformed into z-scores and further explored using box plots to check for outliers in the data. Nine scores were detected as potential outliers in the Perceptual Quality (n = 5), Narrative Consistency (n = 3), and Plausibility (n = 1) subscales, indicated by $2.58 < z < -2.58$ values. Scores for the Plausibility and Perceptual Quality subscales were found to be consistent with participants' responses for the rest of the items in the subscales and were therefore not removed. One of the scores in the Narrative Consistency subscale was, however, considerably lower than the participant's mean score for the subscale (difference of 3.5) and was therefore removed. Examination of Z_s and Z_k scores, and Q-Q plots indicated that scores were positively skewed for most groups. Shapiro-Wilks test of normality was significant for three groups (see Table J1). The perceptual quality subscale was skewed and leptokurtic (indicated by z scores larger than $\pm$ 1.96) for one level of the independent variable. While the *t*-test is generally considered to be 'robust' to violations of normality, especially for relatively large sample sizes ($N > 30$), where the groups are not similarly skewed (or distributed), the *t*-test becomes less robust to deviations in normality. Given that this is the case for the perceptual quality (see Table J1), a non-parametric approach to the analysis was taken.

Table J1

Normal distribution statistics for each subscale and condition for Pilot Study 2

Condition		*W*	df	*p*	Z_s	Z_k
Plausibility	Fictional	.96	35	.202	- 1.40	0.53
	Non-fictional	.98	30	.759	- 0.43	- 0.39
Typicality	Fictional	.94	35	.041	1.61	- 0.42
	Non-fictional	.91	30	.012	1.49	- 1.06
Factuality	Fictional	.96	35	.233	0.21	- 1.14
	Non-fictional	.97	30	.535	0.81	- 0.05
Perceptual quality	Fictional	.88	35	.001	- 3.35*	2.29*
	Non-fictional	.95	30	.121	- 1.64	0.16
Narrative consistency	Fictional	.97	35	.344	- 0.91	0.18
	Non-fictional	.93	30	.320	- 1.20	0.03

Note. Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis.

*Indicates skewness and kurtosis *z* scores larger than ± 1.96 and therefore significantly non-normal.

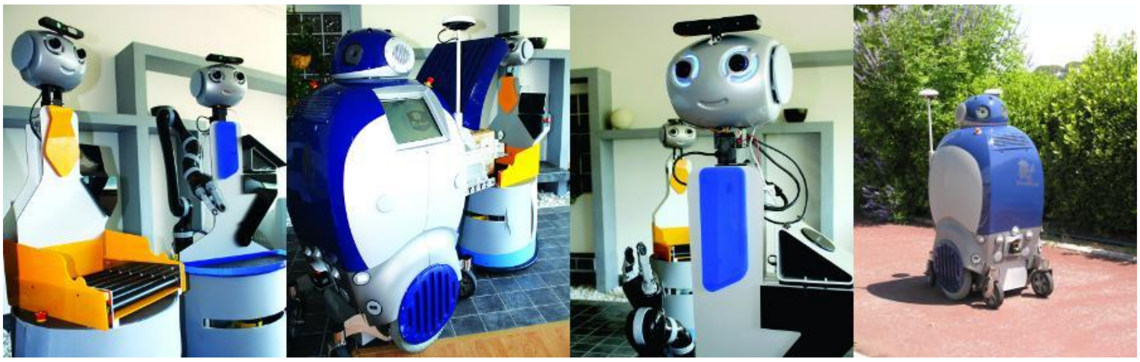
Appendix L

Assumption checks relevant to the analysis presented in the Results section of Pilot Study 3.

Shapiro-Wilk test of normality was significant for all but one of the dependant variables in both conditions (see Table K1) indicating a non-normal distribution. Further examination of Z_s and Z_k values showed that eight variables were substantially negatively skewed (indicated by $2.58 < z < -2.58$ for $N > 50$). Visual inspection of Q-Q plots confirmed this and also indicated that transforming the data was unlikely to correct the distribution as not all variables were skewed in the same way. Given that the one-way MANOVA is considered fairly robust to deviations from normality for relatively large samples, it was decided to proceed with the analysis without any transformation of the data. However, it should be noted that multivariate normality cannot be assumed given that univariate normality has been violated.

Appendix M

Images of Robot-Era robots (from left to right) CORO, DORO, and ORO presented to participants prior to collection of their responses to the anxiety, attitudes toward the use of robots, and perceived enjoyment subscales of the Almere Model. Robots developed as part of the Robot-Era Project (2014).



Appendix N

Supplementary post-hoc analyses carried out for Study 5.

Approximately half or more of the participants believed that video that they watched and the robots depicted within were fictional or non-fictional in accordance with the condition they were assigned to (see Table M1 and M2). This implies that, in general, participants believed the information they were given about the video and the robots. However, it should be noted that a substantial number of participants provided a different response to the one matching their condition assignment. In order to investigate whether the effect of the independent variable on the seven dependent variables was affected by whether participants' believed the manipulation or not (i.e., believed that the robots and story presented in the video were fictional or non-fictional depending on the condition to which they were assigned), separate MANOVAs were carried out with participants who believed the manipulation, and those who did not.

Participants' belief that the robots in the video were fictional or non-fictional

The data was split into two: participants who believed that the robots' fictionality (i.e., fictional or non-fictional) matched the condition to which they were assigned or participants who believed that the robots' fictionality did not match the condition to which they were assigned (see Table M1). Two MANOVAs were conducted for each part of the data set for six of the dependant variables. D-scores from the IAT were analysed separately using two independent samples *t*-test. A Bonferroni correction was applied to account for multiple post-hoc comparisons, resulting in a new critical value of $p = .007$ (7 comparisons) for each MANOVA.

Table M1

Number of participants who believed the robots in the video to be fictional or non-fictional by condition

Condition	"I believed the robots to be…"		
	Fictional	Non-fictional	*N*
Fictional	30 (46.9%)	34 (53.1%)	64
Non-fictional	21 (29.2%)	51 (70.8%)	72

Condition assignement had a significant effect on participants attitude toward robots, Pillai's trace, $V = 0.25$, $F(6, 74) = 4.08$, $p = .001$, for the participants who believed that the robots' fictionality matched the information they were given. The significant MANOVA was followed up with multiple pairwise tests and *p* values are reported in Figure M1 and Figure M2. There was no significant difference of participants d-scores between the two conditions (see Figure M3). The results from this analyses were almost the same as the results from the original MANOVA with one notable exception, there was no significant difference between the conditions for the third subscale of the NARS (emotion in interaction with robots).

Condition assignement had no significant effect on participants attitude toward robots, Pillai's trace, $V = 0.83$, $F(6, 48) = 0.72$, $p = .635$, for the participants who did not believe that the robots' fictionality matched the information they were given. As the MANOVA was non-significant, it was not followed up further (see Figure M1 and Figure M2). There was no significant difference between participants' IAT scores for the two conditions (see Figure M3). This result was inconsistent with the original analysis and may indicate that participants who did not believe that the fictionality (fictional or non-fictional) of the robots in the video matched the condtion to which participants were assigned (fictional or non-fictional) were not significanlty affected by the experiemental manipulation.

Participants' belief that the story in the video were fictional or non-fictional

The data was split into two: participants who believed that the story's fictionality (i.e., fictional or non-fictional) matched the condition to which they were assigned or participants who believed that the story's fictionality did not match the condition to which they were assigned (see Table M2). Two MANOVAs were conducted for each part of the data set for six of the dependant variables. D-scores from the IAT were analysed separately using two independent samples *t*-test. A Bonferroni correction was applied to account for multiple post-hoc comparisons, resulting in a new critical value of $p = .007$ (7 comparisons) for each MANOVA.

Table M2

Number of participants who believed the story in the video to be fictional or non-fictional by condition

Condition	"I believed the robots to be…"		
	Fictional	Non-fictional	*N*
Fictional	41 (64.1%)	23 (35.9%)	64
Non-fictional	38 (52.8%)	34 (47.2%)	72

Condition assignement had a significant effect on participants attitude toward robots, Pillai's trace, $V = 0.20$, $F(6, 68) = 2.81$, $p = .017$, for the participants who believed that the robot's fictionality matched the information they were given. The significant MANOVA was followed up with multiple pairwise tests and *p* values are reported in Figure M4 and Figure M5. There was no significant difference of participants d-scores between the two conditions (see Figure M6). The results from this analyses were almost the same as the results from the original MANOVA with one notable exception, there was no significant difference between the conditions for the third subscale of the NARS (emotion in interaction with robots). Similar to the result described in the previous section, the results from this analyses were almost the same as the results from the original MANOVA with one notable exception, there was no significant difference between the conditions for the third subscale of the NARS (emotion in interaction with robots).

Condition assignement had no significant effect on participants attitude toward robots, Pillai's trace, $V = 0.09$, $F(6, 54) = 0.87$, $p = .523$, for the participants who did not believe that the story's fictionality matched the information they were given. As the MANOVA was non-significant, it was not followed up further (see Figure M4 and Figure

M5). There was no significant difference between participants IAT scores for the two conditions (see Figure M6). This result was inconsistent with the original analysis and may indicate that participants who did not believe that the fictionality (fictional or non-fictional) of the story in the video matched the condtion to which participants were assigned (fictional or non-fictional) were not significanlty affected by the experiemental manipulation.

Figure M1

Mean anxiety, attitude, and perceived enjoyment ratings for both conditions. Left: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.

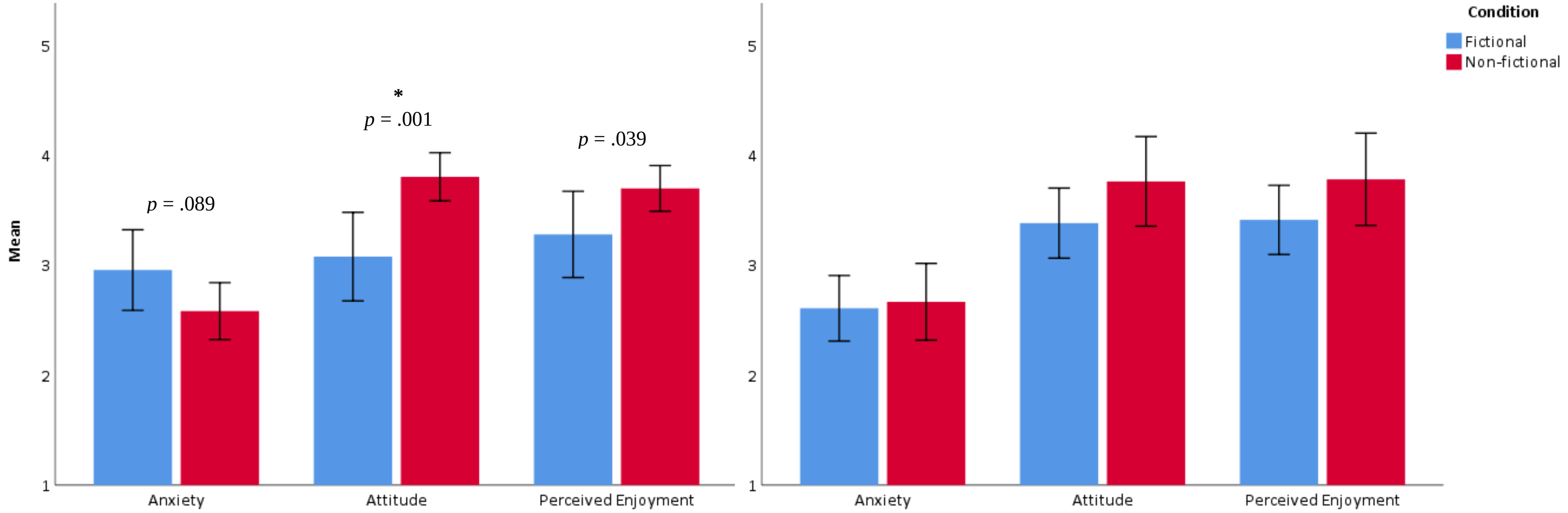


Figure M2

Mean ratings for each subscale of the NARS for both conditions. Left: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.

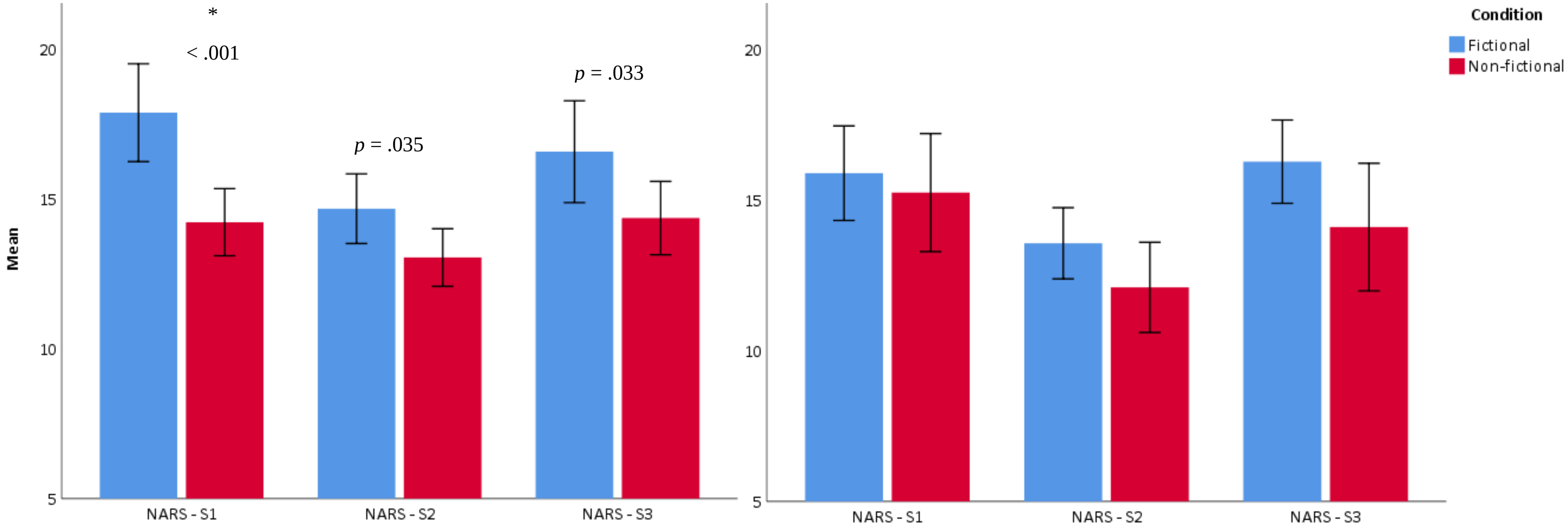


Figure M3

Mean IAT scores for both conditions. Left: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the robots' fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.

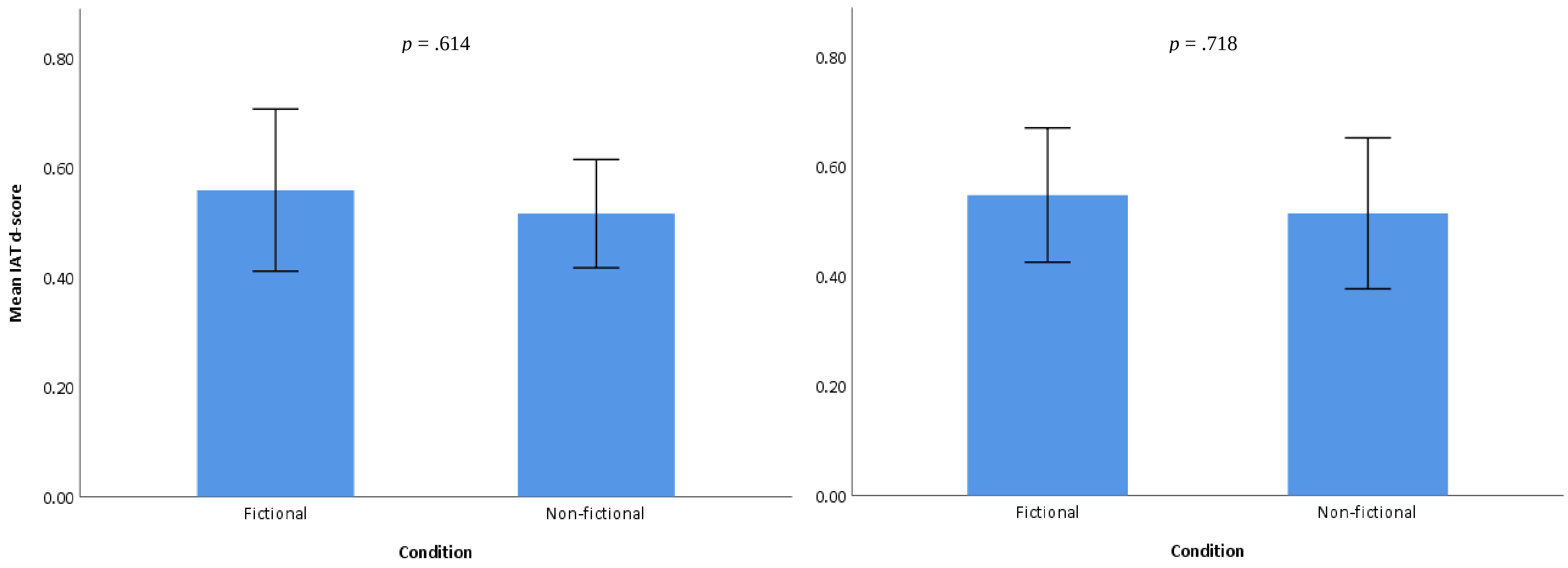


Figure M4

Mean anxiety, attitude, and perceived enjoyment ratings for both conditions. Left: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.

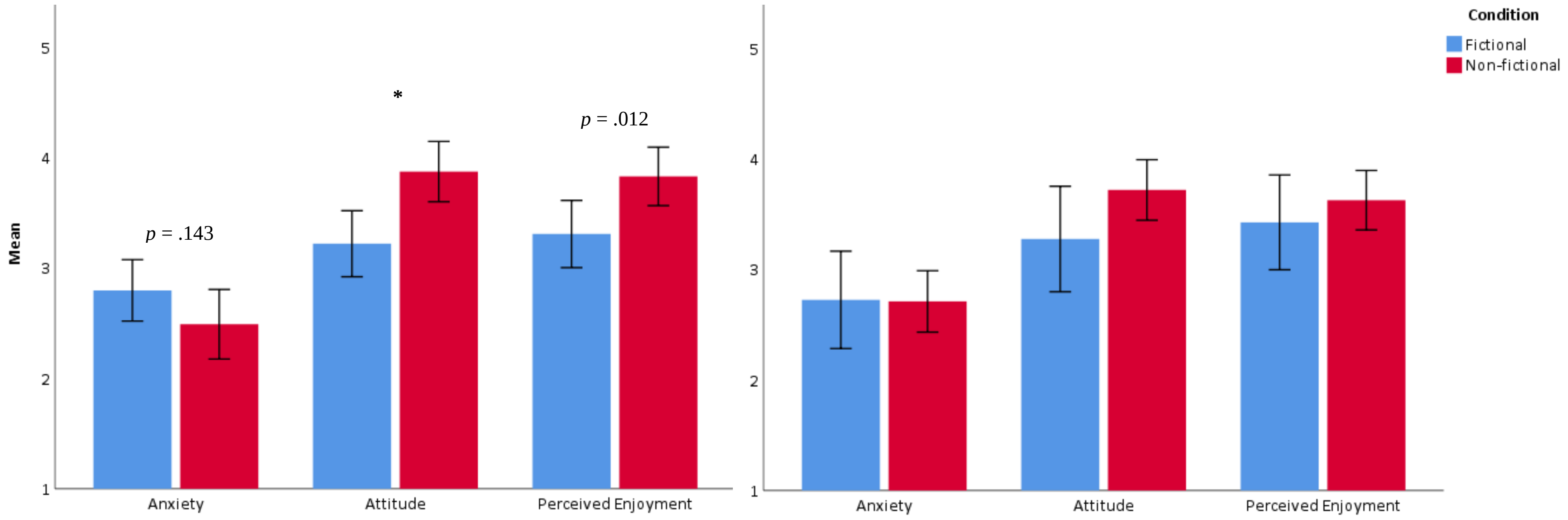


Figure M5

Mean ratings for each subscale of the NARS for both conditions. Left: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.

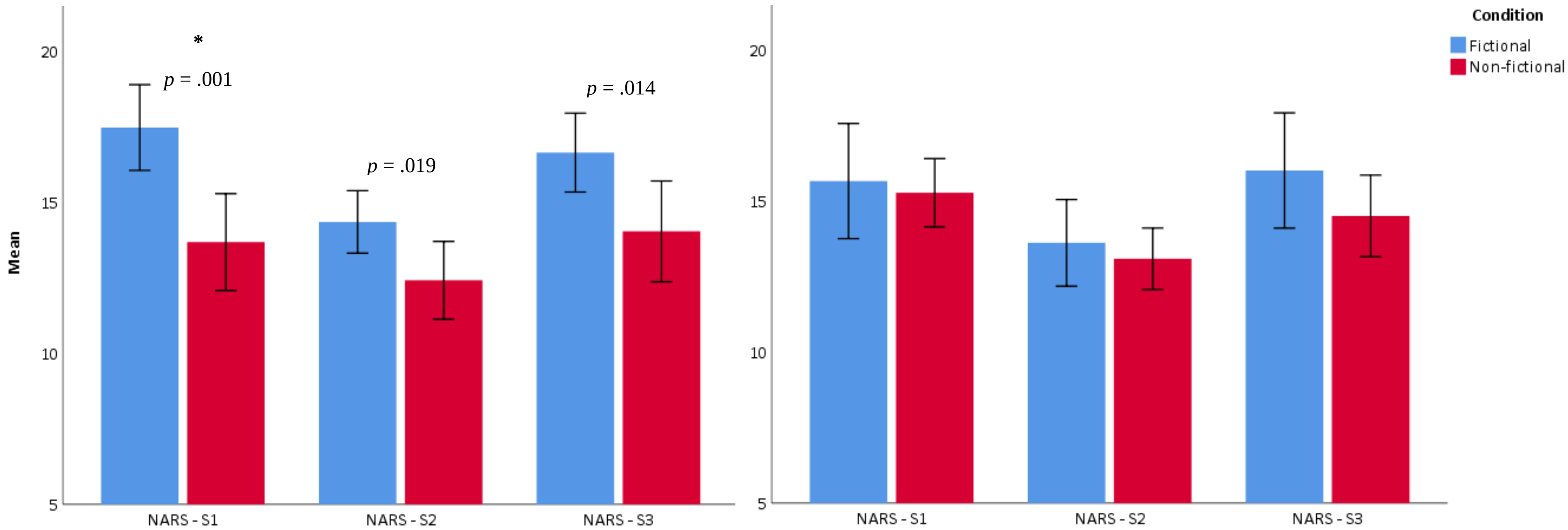
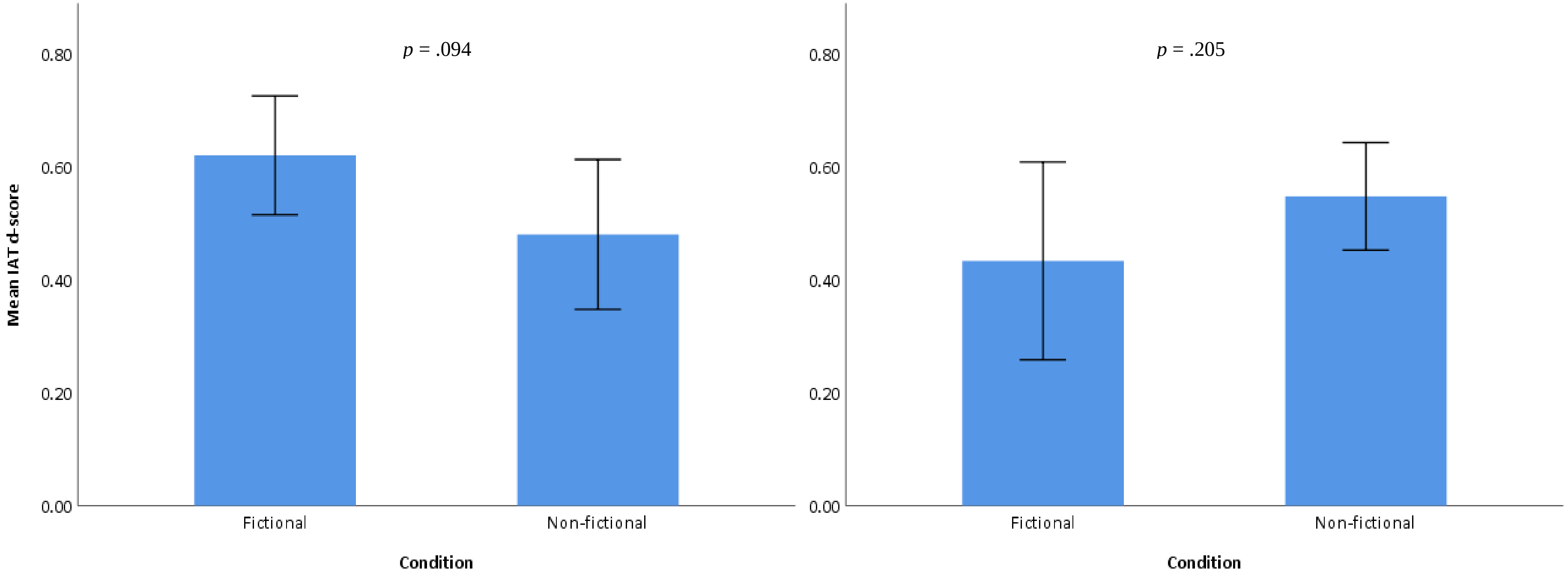


Figure N6

Mean IAT scores for both conditions. Left: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Right: results from the analysis with only participants who believed that the video story's fictionality matched that of the condition to which they were assigned. Error bars are 95% Confidence Intervals.



Appendix O

List of fictional and non-fictional robots that were evaluated in Pilot Study 4.

Fictional Robots

- T-800
- Ultron
- T-X
- NS-5
- Megatron
- C-3PO
- Bumblebee
- Baymax
- Optimus Prime
- RoboCop
- Marvin the Paranoid Android
- Rodney Copperbottom
- Wall-E
- Cyberman

Non-fictional Robots

- Asimo
- Atlas
- T-HR3
- Nao
- Robothespian
- Pepper
- Romeo
- Kengoro
- ARMAR-6
- Zeno
- HUBO
- Valkyrie
- Robina
- Justin

Appendix P

Treatment of the data and assumption checks relevant to the analyses carried out in the Results section of the Pilot Study 4.

In order to check for univariate outliers, participants' responses for each scale and for all 28 images were transformed into *z*-scores and 430 (1.38% of the raw data) potential outliers, as indicated by z-scores of $2.58 < z < -2.58$, were detected (see Table O1). Each potential outlier was checked to see whether it matched the pattern of responses the participant provided for that particular robot. For example, it is unlikely that a participant would rate the robot Baymax as both extremely safe (1) and extremely hostile (1). If this was the case, the response detected as a univariate outlier was deleted. Univariate outliers that were a part of a response set with the same value (e.g., a rating of 1 for each dimension for a specific robot) were not deleted as they were likely to be multivariate outliers and as such would be dealt with in the proceeding step.

To check for multivariate outliers, Mahalanobis distance was calculated for each participant and robot separately, and 117 potential outliers were detected ($p < .001$), indicated by a critical value > 24.32 for seven dependant variables (see Table O1). These potential outliers were not deleted immediately as different patterns of responses were not necessarily an indication of careless responding or response errors/mistakes but could instead reflect an expected variation in responses given the relatively large sample. The responses were scanned to check whether there was in indication of careless responding (e.g., same response for all dimensions) or whether there was a reasonable pattern that would indicate a genuine, albeit unusual, response. Responses that appeared not to adhere to a logical pattern (e.g., robot rated as both extremely safe and extremely hostile) were deleted. From the entire sample of 159 participants, 98.81% of the data was included for further analysis.

Following the exclusions, Shapiro-Wilk test of normality was significant ($p < .05$) for all of the ten robots on all six of the measures (pleasant-unpleasant, hostile-friendly, safe-threatening, familiar-unfamiliar, evil-good, humanlike-mechanical) indicating a non-normal distribution. Visual inspection of Q-Q plots indicated severe skewing for some variables and examination of Z_s and Z_k values confirmed that 28 (approximately

47% of the variables) of the variables were significantly skewed and 16 were significantly kurtotic (indicated by $2.58 < z < -2.58$ for $N > 50$). Distributions of the ratings given to Baymax were particularly problematic with Z_s values raging between 10.07 and -10.53 and Z_k values raging between 15.34 and -2.62. This was not entirely unexpected as, on average, Baymax received the most extreme ratings for all dimensions. It should be noted that the distributions of responses for the non-fictional robots were considerably less non-normal than responses for the fictional robots.

Table O1

Number of detected and deleted univariate and multivariate outliers for each robot

		N of detected univariate outliers	*N* of deleted univariate outliers	*N* of detected multivariate outliers	*N* of deleted multivariate outliers *
Fictional Robots	Baymax	41	12	10	4
	BumbleBee	10	-	4	-
	C3P0	22	2	7	3
	Cyberman	12	-	3	3
	Marvin	13	3	-	-
	Megatron	25	4	8	3
	OptimusPrime	2	-	3	2
	Robocop	7	1	6	1
	Rodney	6	-	3	3
	Sonny	33	3	7	4
	T-800	22	3	6	2
	T-X	12	3	2	1
	Ultron	23	5	9	2
	WALL-E	38	2	10	4
	Total	266	37	78	32

* Each multivariate outlier contains seven data points (a rating for each dimension) and therefore means that for each deleted multivariate outlier, seven data points were deleted.

Table O1 (continued)

		N of detected univariate outliers	*N* of deleted univariate outliers	*N* of detected multivariate outliers	*N* of deleted multivariate outliers *
Non-fictional Robots	ARMAR-6	11	-	2	2
	Asimo	29	3	1	1
	Atlas	8	-	1	-
	HUBO	21	2	3	1
	Justin	13	-	4	1
	Kengoro	2	-	-	-
	NAO	13	3	4	3
	Pepper	16	7	1	1
	Robina	8	1	1	1
	Robothespian	13	-	1	-
	Romeo	6	1	2	1
	THR3	11	-	1	-
	Valkyrie	2	-	1	1
	Zeno	11	3	2	1
	Total	164	20	24	13

* Each multivariate outlier contains seven data points (a rating for each dimension) and therefore means that for each deleted multivariate outlier, seven data points were deleted.

Appendix Q

Treatment of the data and assumption checks relevant to the analyses carried out in the Results section of Study 6.

In order to check for univariate outliers, participants' responses for each measure (for each of the two conditions) were transformed into *z*-scores and 9 potential outliers, indicated by $1.96 < z < -1.96$ values, were found. However, no data points were removed as all identified "outliers" were: a) not extreme outliers, b) in keeping with the individual participant's other scores, and c) unlikely to be a problem for studies with relatively large samples ($N > 50$). To check for multivariate outliers, Mahalanobis distance was calculated for each participant. No multivariate outliers were identified based on Mahalanobis distance critical value > 18.47 for four dependent variables. In conclusion, the assumption of multivariate normality was likely met.

Analysis without covariates

Shapiro-Wilk test of normality was significant for only one the dependent variables in the fictional condition (see Table P1) indicating a non-normal distribution. However, upon further examination of standardised skewness and kurtosis values and Q-Q plots, all dependent variables were deemed to be sufficiently normally distributed and no transformations of the data were performed. Examination of scatterplots for both conditions indicated that there is an approximately linear relationship between the dependent variables. There appeared to be no multicollinearity between variables as indicated by weak to moderate correlations between the dependent variables (see Table 5.25 and Table 5.26 in Section 5.4.2.3). The assumption of homogeneity of variance-covariance matrices was met as indicated by Box's test of equality of covariance matrices ($p = .95$). Variances were homogeneous for all variables as assessed by Levene's Test of Homogeneity of Variance ($p > .05$).

Analysis with covariates

All assumptions relating to the MANCOVA were evaluated and described above. Additionally, the assumption of homogeneity of regression slopes was met, as indicated by a non-significant interaction between evil-good ratings of the robots and the condition, $F(4, 49) = 2.30$, $p = .072$.

Table P1

Distribution statistics for each dependent variable for each condition

Measure	Condition	W	df	$p <$	Z_s	Z_k
IAT	Fictional	0.96	28	.318	-1.19	-0.53
	Non-fictional	0.98	28	.907	0.20	-0.41
NARS-S1	Fictional	0.97	28	.510	0.65	-0.53
	Non-fictional	0.97	28	.528	0.53	-0.53
NARS-S2	Fictional	0.96	28	.291	0.20	-0.94
	Non-fictional	0.94	28	.093	-0.44	-1.41
NARS-S3	Fictional	0.90	28	.014*	1.56	-0.83
	Non-fictional	0.97	28	.540	0.62	-0.64

Note. Z_s is the standardised value of skewness, and Z_k is the standardised value of kurtosis

* indicates substantial departure from the normal distribution