Towards Robot-Assisted Therapy: Identifying Mechanisms of Effect in Human-Biomimetic Robot Interaction

Emily C. Collins

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

September 2016

Gloria had a grip about the robot's neck that would have asphyxiated any creature but one of metal. Robbie's chrome-steel arms wound about the little girl gently and lovingly, and his eyes glowed a deep, deep red.

"Well," said Mrs Weston, at last, "I guess he can stay with us until he rusts."

Robbie, Isaac Asimov, September 1940

Acknowledgements

I would like to thank the many people who have advised, encouraged and supported me throughout my time writing this thesis. I am particularly indebted to the following.

My supervisor, Tony Prescott, for everything. Our conversations have always inspired and challenged me, may they always continue to do so. Ben Mitchinson, for introducing me to Robbie, and like Robbie for Gloria, being there for me and helping me stay out of harm's way. Jonathan Aitken, David Cameron, James Law, Owen McAree, Noel Sharkey and Michael Szollosy, for not least the many amusing and caring corridor conversations. Sam Fernando, Abigail Millings, Roger Moore, Amanda Sharkey and all the other members of the EASEL Project, for helping me to expand my knowledge. The remaining members of Sheffield Robotics, for providing us all with a wonderful place in which to do research. Hiroshi Ishiguro, Yoichiro Yoshikawa, Takamasa Iio, Shogo Nishiguchi and everyone at the Intelligent Robotics Laboratory, Osaka University. Sebastian Conran, for introducing me to the world of industrial design.

I would like to thank all my friends who have shared so much with me since I came to Sheffield. I am particularly grateful to Kirsten Bartlett, Iveta Eimontaite, Lizzy Kirkham, Connor Little, Michelle McGillion, Leila and Rosie Prescott, Robin Scaiff, and Dan Wilson. And, of course, The Brain Trust: Craig Bertram, Ben Dornan and Maria Panagiotidi: One For All and All For Obscure Music References.

Finally, I would like to thank the Engineering and Physical Sciences Research Council, The Japan Society for the Promotion of Science, the Framework 7 project EASEL (FP7-ICT-611971), and The University Of Sheffield, for the financial support I have received.

Declaration

I hereby declare that all the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Some of the work reported in this thesis has been published elsewhere:

Collins, E. C., Millings, A., & Prescott, T. J. (2013). Attachment to Assistive Technology: A New Conceptualisation. In *Proceedings of the 12th European AAATE Conference (Association for the Advancement of Assistive Technology in Europe).*

Collins, E. C., Millings, A., & Prescott, T. J. (2014). Individual Differences and Biohybrid Societies. In *Biomimetic and Biohybrid Systems* (pp. 374-376). Springer International Publishing.

Collins, E. C., Prescott, T. J., & Mitchinson, B. (2015). Saying It with Light: A Pilot Study of Affective Communication Using the MIRO Robot. In *Biomimetic and Biohybrid Systems* (pp. 243-255). Springer International Publishing.

Collins, E. C., Prescott, T. J., Mitchinson, B., & Conran, S. (2015). MIRO: A Versatile Biomimetic Edutainment Robot. In *Proceedings of the 15th Conference on Advances in Computer Entertainment Technology.* ACM, New York, NY, USA.

Millings, A. & Collins, E. C. (In press). Human Relationships with Living

Machines. In *Living Machines: A Handbook of Research in Biomimetic and Biohybrid Systems*. Prescott, T.J. and Verschure, P.M.J (eds), Oxford University Press.

Emily C. Collins, 1 September 2016.

Thesis Summary

This thesis provides a framework for understanding human-robot relationships based on human-other bonds. It focuses on human-animal interactions and the positive effects of Animal-Assisted Therapy (AAT), proposing that Robot-Assisted Therapy (RAT), with biomimetic robots, could benefit from a better understanding of the mechanisms of effect driving positive AAT outcomes. In sum, interactions with biomimetic robots could provide benefits mechanistically comparable to those provided by AAT animals, independent of individual differences in personality or culture.

Evidence is provided showing that interaction with a PARO therapeutic robot led to a positive change in users' well-being, measured via Felt Security (FS). Intimate interactions with PARO, such as stroking the unit, produced greater increases in user FS, independent of individual differences in caregiving and attachment styles.

The Felt Security Scale (FSS) was translated into Japanese creating the JFSS. This was used in a cross-cultural study (Japan/UK) which demonstrated that the biomimetic robot MIRO does not have to display predictable behaviour in order to have a positive impact on a user's FS. These results were found in both the UK and Japan despite the different culturally-driven expectations of robot-acceptance in the two countries.

Although an interaction with both PARO and MIRO increased user FS, these scores were significantly higher when interacting with PARO.

Contents

Chapter 1

Introduction and Overview

1.1 Summary

This thesis is an investigation of some of the open questions about how human-biomimetic robot interactions can positively impact a user's well-being. The approach is two-fold, reflective of the human-robot dyad. Firstly, the investigation explores the individual differences of the human, to enquire as to whether the tailoring of a robot to a user has the potential to lead to improvements in a robot's performance as measured by its impact upon the user's psychology. Secondly, the investigation asks how specific mechanisms of the robot itself affect its performance as measured by its impact upon the user. To achieve this a comparative approach is taken. As this thesis is focussed on biomimetic robots, the domain within which this thesis lies is that of human-animal relationships. To explore this domain, tooling from existing areas of research in human-other relationships are applied, with specific focus on the individual differences of users which may impact the outcome of a human-robot interaction.

This chapter provides the motivation for this work by introducing the concept of human-robot relationships. It concludes with an outline of the research objectives, and a detailed summary of the content of the subsequent chapters.

1.2 The Development of Human-Robot Interactions

The term ‘Robotics’ was introduced by Issac Asimov at the beginning of the twentieth century.[1] From his earliest works of science-fiction the themes addressed by Asimov can be interpreted as markers for modern expectations in robotics. For example, in September 1940 Asimov published *Robbie*, the first story in his popular science *Robot* series (Asimov, 1940).[2] In *Robbie*, the eponymous robot belongs to the Weston family, serving as a nursemaid for their daughter, Gloria. With the rise of anti-robot sentiment, Mrs. Weston decides Robbie must go. She fears how attached her daughter has become to the robot, and witnesses Gloria shunning other playmates in favour of spending her time with Robbie. Mr. Weston concedes and the unit is returned to its factory. Gloria is distraught, and in spite of her parents buying her a dog her condition deteriorates. Believing that Gloria cannot forget Robbie, because she thinks of the robot as a person and not a machine, her parents take her on a tour of a robot construction factory. During the tour they come across Robbie, working

---

[1] The word ‘Robot’ itself was first used to denote artificial biological organisims, in Karel Čapek’s 1920s play *R.U.R. (Rossum’s Universal Robots)*. The word had been suggested to Karel by his brother Josef, as a replacement for the word *laboři* (‘labour’ or ‘workers’; Zunt (2002)). ‘Robot’, or *robota*, in Čapek’s native Czech, means *corvée*, the unpaid work period a vassal owes their lord. After this first use by Čapek, the word began to displace ‘automaton’ and ‘android’ in languages around the world. Its use in popular science publications, such as those by Issac Asimov, spread and cemented the term as denoting a machine capable of carrying out complex actions automatically.

[2] Asimov’s *Robot* series features machines with fictional positronic brains, which function as central processing units and provide the robots with a form of consciousness whilst being hardwired with the Three Laws of Robotics.

as an assembler. In her joy at finding her friend, Gloria runs out onto the factory floor in front of a moving vehicle, whereupon Robbie rescues her. Mrs. Weston, in seeing Robbie save Gloria's life, decides the robot is not a soulless monster and accedes: "Well, I guess he can stay with us until he rusts" (Asimov, 1940).

Nearly eight decades later, these themes are not science-fiction, but serious considerations for robotic's researchers. Not only is the concept of a robot companion like Robbie, which can compete with a dog for a human's affection, science fact, but these themes of anti-robot sentiment, fears of detrimental attachment to machines, and the confusion by a user that a robot is neither alive nor inanimate, frame the current climate of social robotics.

On May 11th, 1999, Sony released its first consumer model of AIBO: Artificial Intelligent Robot (though *aibō* also means 'pal' or 'partner' in Japanese). AIBO is a robotic pet marketed for domestic entertainment use. In 2006, Sony discontinued AIBO, and from July, 2014, Sony stopped servicing them (Wikipedia, 2016). Now, to have an AIBO serviced many owners go to the home workshop of Nobuyuki Norimatsu, a former Sony employee. Norimatsu took over the servicing of AIBO's upon noticing that many owners felt sad when their AIBO broke down. He explains that for people who otherwise cannot have pets, such as restaurant owners or those with allergies, the AIBO serves as a replacement and becomes like a member of the family (VProBacklight, 2016).

Figure 1.1. PARO the therapeutic robot, developed by AIST in Japan, is an example of a commercially available social biomimetic robot.

In a world where robots as are ubiquitous as they are in 2016, the idea of a robot becoming akin to a family member is not unusual. Robots, that is embodied artificial agents, are found in industry, the service sector, space exploration, the military, mining, healthcare, and the home (though this list is not exhaustive). The focus of this thesis are those robots found in these last two categories. Social robots are here defined as artificial agents that have a physical presence, some autonomy and the ability to adapt and communicate (Collins et al., 2013).

Social robots are increasing in functionality and use in healthcare and home environments (Collins et al., 2013). The recent commercial success of PARO the therapeutic robot (Shibata, 2016) is indicative of a real need in at least the healthcare market for these advanced tools (See Figure 1.1).

The focus of this thesis is specifically on social biomimetic robots. A biomimetic machine is one whose design is inspired by nature, via an imitation of the models, systems, and elements of nature for the purpose of solving human problems (Pearson et al., 2011). Robots built to resemble and behave like animals with the purpose of serving a role as a companion fall firmly into this definition.

1.2.1 The dawn of human-robot relationships

How technology is used depends on how a person relates to it. As technologies such as social robots are developed, the relationships that humans have with them will become more complex, and may take on some of the characteristics of the social relationships that humans have with other living things. This is important to note. Such relationships have the potential to be either beneficial or harmful to the long-term welfare of the robot user. UN global demographic's projections from 2012 indicate a continued increase in the global population in the near future with a decline in population growth, with the global population expected to reach between 8.3 and 10.9 billion by 2050 (UN, 2012). The increase in population rates coupled with the decrease in population growth indicate a disparity between global populations of working adults against retirees. One solution to the global ageing population then, is an increase in assistive living companions and other assistive technologies providing the elderly with the means to live independently. That requires the development of sophisticated robotics capable of handling daily assistive tasks, whilst simultaneously safeguarding the well-being of individuals by investing in research that explores the impact of potential social relationships with robotic agents.

Furthering understanding of human-robot relationships entails both learning what it is about a user that directs how they relate to a robot, and also learning what it is specifically about robots the influences their user. This information would lead to more user appropriate and subsequently efficient devices.

1.2.2 Societal implications of human-robot relationships

Ethical issues have been raised as a result of the expected increase in accessible robotic technology intended for consumer use (Sharkey and Sharkey, 2010). Established interactive technologies, most notably those Internet-enabled, are already having a profound impact on society. The Internet, previously the domain of a static device kept in a single room, is now available via a proliferation of enabled devices such as laptops, mobile phones, and games console. Its accessibility has significantly changed how individuals engage with the world. For example, via the Internet adolescents can easily encounter and consume sexually explicit material (Owens et al., 2012). The impact of Internet pornography on sexual attitudes, beliefs, behaviours and sexual aggression is of great interest to researchers and policy makers. Self-concepts, body image, social development, brain function and physical development are all under scrutiny through the lens of the Internet, humanity's increasing Internet addiction (Kuss et al., 2014), and the transformative effect of the simple act of Smartphone ownership upon culture (Sarwar and Soomro, 2013). There is no reason to expect that the societal impact of social robots will be any less profound.

Concerns have already been raised that social robots in the form of assisted living companions will lead to an increase in loneliness (Sharkey and Sharkey, 2010), or to a decrease in social responsibility from individuals content to allow a robot to take the place of a human carer (Sparrow and Sparrow, 2006). Whilst an alternative view is that the increase in robotic technology to provide the elderly with a means of independence will in fact improve social interactions, as the technology frees individuals to engage more easily with the environment beyond their home. A parallel can be drawn between the Internet and social robotics in this case.

In 2003, Amichai-Hamburger and Ben-Artzi noted that heavy Internet users seemed alienated from normal social contacts (Amichai-Hamburger and Ben-Artzi, 2003), as had been discussed by Kraut and colleagues in 1998 who directly argued that Internet use led to loneliness (Kraut et al., 1998). However, despite the association between Internet use and increased loneliness via reported declines in face-to-face communication, it has been alternatively reported that lonely individuals use the Internet as a means of alleviating their loneliness. Morahan-Martin and Schumacher (2003), reported that lonely individuals were more likely to use the Internet for emotional support than individuals who self-reported as not-lonely. The anonymity and lack of face-to-face communication offered by online interactions helped this sub-set of individuals to achieve satisfactory social behaviour and modulate existing negative moods (Morahan-Martin and Schumacher, 2003). This demonstrates that there are always pros and cons to be considered when societal changes are being witnessed in the wake of a technological paradigm shift. In the case of the coming influx of commercially available social robotics societal changes will be large. However, these changes can also be expected to take a different form to those that have occurred as a result of the Internet.

As defined above a robot is an embodied artificial intelligence, with some degree of autonomy, and, in the case of a social robot, communicative ability. The Internet is a multimedia platform that offers anonymity to its user. Internet-enabled devices are directly used by humans to interact with the world. A social robot is different. It is tangible and independent. Robots remain *laboři*: they work for humans as tools which extend a person's presence and abilities, but as the technology advances the relationships that humans are capable of having with robots are evolving.[3] A social robot's physical presence, in particular a social biomimetic robot, is purposed as a machine made to imitate a living creature. Robots are not alive. They are not other humans or other animals, but neither are they merely objects. Their ontological status might be best described as 'liminal' (Kang, 2011). That is neither living as a biological entity is, nor purely mechanical as with a traditional machine, and a robot's liminal categorisation is indicative of its unknown future role as one half of a human-robot

---

[3] *Laboři* translates from Czech to 'labour' or 'workers' in English. It was the original word Karel Čapek used in his play *R.U.R. (Rossum's Universal Robots)* to describe the mechanical beings. *Laboři* was subsequently replaced with *robota* ('robot').

relationship.

Anecdotal evidence suggests that the relationships that humans are already forming with their robot tools are as strong as they might be with living things (e.g., Carpenter (2016); Friedman et al. (2003); Sung et al. (2007)). The use functions that robots are built to serve include those with industrial applications, but more frequently are robots being built to support a human's emotional and/or social needs (for example see the use of interactive robots in autism therapy, e.g., Robins et al. (2005)). This development begs a serious question: how are researchers to analyse these purposefully social relationships? How are we to describe what a human-robot relationship looks like in all its forms, such that we might know when a relationship is becoming damaging, or when it is likely to produce positive outcomes?

In sum, due to the expected increase in social robot use, and the associated societal and legal ramifications, it is prudent to develop a consistent framework with which to describe and analyse human-robot relationships. A framework that takes into consideration how humans relate to their existing environments.

1.2.3 A new approach to understanding human-robot relationships

There has already been some attempt to explore human-robot interaction (HRI) from a human factors perspective, and this will be discussed in Chapter 3. However, the new approach proposed in this thesis for understanding human-robot relationships is comparative. Rather than considering the robot as an established entity taking up one half of the human-robot dyad, and exploring that robot in and of itself, the work presented in this thesis considers the robot half of the dyad against a backdrop of other agents with whom, or with which, a human may form a relationship. In this manner the human is central to the analysis.

Though the contributing aspects of both the robot and the human are considered independently, what is known about how a human develops a relationship with other agents forms the foundation of how best to analyse the impact of any particular robot upon an individual. In doing this the work included in this thesis attempts to demonstrate the importance of breaking down the specifics of a human-robot interaction to see what aspects of the dynamic are leading to what outcomes.

This thesis demonstrates that a comparative approach is possible. It provides examples of how to approach an understanding of the influential aspects of a human-robot interaction. What is it about the person, or the robot specifically that contributes to a positive outcome? In knowing the answers to these questions, and in having a demonstrable framework in which to ask and explore these questions, future robots can be built which provide the most efficient outcome for a user.

1.3 Content of the Thesis

The motivation for the work described in this thesis has been explained above via the introduction of the concept of a human-robot relationship. There are many open questions left in HRI. This thesis has been inspired specifically by the question of what it is about each individual HRI case that results in positive outcomes, with a particular emphasis on understanding the case of social biomimetic robot applications, which are those most likely to evoke the conditions for a relationship. In seeking to understand this a framework is required that takes the mechanism and performance of a robot into consideration, as well as the individual nature of the user. It is these ideas which constitute the focus of the thesis.

The first objective, which is described in Chapters 2 and 3, is the establishment of the domain and tooling that underlie the experimental work. This culminates at the end of Chapter 3 with the introduction of a new framework of human-robot relationships built upon existing understandings of human-other relationships.

The second objective, dealt with in Chapter 4, is prompted by this framework, and considers to what extent the individual differences of the human user influence the outcome measure of a human-robot interaction. This is based on the consideration that caregiving or attachment style may influence how an individual bonds with other humans, animals and objects. An idea introduced in Chapter 3 via the framework of human-other relationships.

The third objective, which is the topic of Chapter 7 deals with the precise mechanisms of the robot that may influence the outcome measure of a human-robot interaction. Specifically it explores the extent to which the predictable mammalian-like behaviour of a biomimetic robot is required to produce a positive emotional response in a user. This question is vital to the efficient use of generic animal-like robots for therapy, in that the answer allows developers to understand how sophisticated a robot needs to be in order to be a useful health care tool. Chapter 5 begins to address this objective by presenting MIRO, a robot platform that is suitable for systematic investigation of mechanisms of effect in RAT, as well as the experimental validation of one aspect of its design. Chapter 6 returns to the issue of a user's individual differences by exploring the impact of a person's culture on the measurement of a robot's performance. To that end Chapter 6 discusses cross-cultural differences in robot acceptance, and presents the translation of the Felt Security Scale (FSS) into Japanese. Chapter 7 utilises both MIRO and the Japanese FSS, and presents a cross-cultural study which explores the impact that predictable mammalian-like behaviour has on felt security in both the UK and Japan.

Finally, Chapter 8 offers some conclusions drawn from the work in this thesis and considers future research directions.

The following is a detailed summary of the subsequent chapters:

Chapter 2 - AAT and RAT: Animal and Robot Assisted Therapies begins with a review of the Animal-Assisted Therapy (AAT) literature that forms the theoretical domain of this thesis. The impact of animals on therapy is discussed, and used to introduce the use of robots in therapy (Robot-Assisted Therapy (RAT)). In particular the design of therapeutic robots prompts a discussion of the mechanisms at play in both AAT and RAT. In an absence of clarity on mechanism in either domain the chapter calls for more controlled studies that explore what it is specifically about robots that aids positive changes in RAT users. The chapter concludes by asking what metrics, psychological and otherwise, are available for measuring performance in AAT that might be relevant to RAT analysis.

Chapter 3 - Modelling Human-Other Relationships introduces a framework for analysing and describing human-robot relationships based on existing research in human-human, -animal and -object bonds. This chapter considers the tooling that can be used to explore theory and methodology in HRI via comparing it to tooling that is already applied in established human-other research.

Chapter 4 - PARO: Felt Security, Physical Engagement and Individual Differences. Applying the framework developed in Chapter 3, Chapter 4 takes as its starting point results from human-other literature which indicate that individual caregiving or attachment styles lead to varying approaches in human-other interactions. Chapter 4 asks: Does a person's caregiving or attachment style impact the outcome of a human-robot interaction, as it does with other human relationships, in particular the human-animal relationships seen in AAT? In AAT the attachment style[4] of the human is sometimes given consideration when developing suitable AAT interventions, based on assumptions that a person's attachment style has an influence over how that person interacts with others in their environment. A metric of change in emotional state, measured as felt security delta, was taken. Felt security was chosen because it is associated with caregiving and attachment styles in the human-human interaction literature. Felt security has been conceptualised by past research as comprising feelings of care, esteem, love and safety. These concepts are cornerstones to AAT, where patients include individuals who have suffered as a result of a lack of safety and security in their lives. Felt security is also linked to feelings of loneliness, another area that AAT and RAT are used to treat. A coding scheme was developed over the course of a pilot to main study. This coding scheme of engagement through touch allowed quantitative description of the interaction. Despite promising results obtained in the pilot study that indicated caregiving and attachment style might be indicative of an individual's physical engagement behaviour, the main study revealed

---

[4] Caregiving style has not yet been explored in AAT.

that caregiving and attachment styles were not predictive of observed physical engagement with the robot. However, after controlling for caregiving and attachment styles, felt security did increase after interacting with the PARO robot. Further, how the individual physically engaged with PARO during the interaction was related to felt security delta, such that individuals who spent more of their time engaging PARO in touches that were coded as *intimate* had greater felt security delta scores than individuals who spent most of their interaction time not touching PARO at all. Thus greater emotional engagement (EE) does appear to be mechanistic in improving felt security during an interaction with a robot intended for RAT. Chapter 4 demonstrates that human-other methods can be useful in exploring HRI.

Chapter 5 - MIRO: A Biomimetic Platform for Systematic Investigation of RAT. Chapter 4 demonstrated one way in which RAT is effective at leading to a positive psychological change in a user. However, PARO, the platform used in Chapter 4, is a robot that displays only simply behaviours. For an experiment which aims to discover mechanisms of effect researchers require a robot that can be selectively modified in order to test for features useful for RAT. Therefore, in Chapter 5 the selectively controllable biomimetic platform, MIRO, is introduced. The chapter also presents an experiment validating one aspect of MIRO's design: its ability to communicate affect via light. The chapter highlights the platform's suitability for conducting controlled HRI studies.

Chapter 6 - Cultural Effects and Translating the Felt Security Scale. Chapter 6 discusses the effect of culture on acceptance and believability in HRI, and presents the translation of the Felt Security Scale in to Japanese to produce the new JFSS instrument. Further, the production of this new instrument helps to address the question of how transferable the comparative methodological approach is to cross-cultural work.

Chapter 7 - MIRO in the UK and Japan: Felt Security and Behaviour. Chapter 4 indicated that a social biomimetic robot can impact an individual's felt security, and that this is related to the level of physical engagement an individual has with the robot during an interaction with it. Inspired by this the work presented in Chapter 7 was designed to investigate what it is about a social biomimetic robot that impacts felt security in a user. Employing the framework developed in Chapter 3, and its description of a gradation of interaction between a human and another that exists on a continuum of living to inanimate, as well as the AAT domain of this thesis, the focus of Chapter 7 is on understanding the predictable mammalian-like behavioural mechanisms at play during a human-robot interaction.

As Chapter 4 indicated that emotional engagement (EE), as measured via physical engagement, is important to the level of a user's felt security delta score, Chapter 7 is an exploration of whether predictable behaviour, an aspect thought to be important in engaging a user, was related to an increase in felt security. Chapter 7 asks: Is it the predictable mammalian behaviour of a social biomimetic robot that leads to an impact on a user's felt security? Further, given existing literature which proposes that the Japanese respond to robots in a markedly different manner to the British (as discussed in Chapter 6) Chapter 7 also asked: Does the impact a robot has on its user depend on the user's cultural background? A controlled cross-cultural study was designed, to be conducted in both Japan and the UK, in which interactions with a robot displaying either predictable or incongruent behaviour were measured. No difference was found between groups. Chapter 7 begins to offer answers as to what it is about a social biomimetic robot that impacts felt security in a user. In this case it does not appear to be related to consistency in predictive mammalian behaviour. Further, despite literature proposing the existence of cross-cultural differences in responses to robots, the study showed no difference in how Japanese or British participants responded to a robot that displayed either predictably mammalian or incongruent behaviour.

Chapter 8 - Conclusions and Future Work is a summary of the work included in this thesis. Chapter 8 considers what advances have been made overall in the field of HRI by this work, and proposes several avenues for future research continuing on from the findings of the thesis.

1.3.1 A note on collaborative work in this thesis

Some work presented in this thesis was completed in collaboration with other members of Sheffield Robotics at The University Of Sheffield. This is reflected in the collaborative papers listed in the Declaration (p. iii).

MIRO described in Chapter 5 was designed by Sebastian Conran Associates and built by Eagelmoss Ltd., with control architecture designed and built by Dr. Ben Mitchinson of Sheffield Robotics.

Work included in Chapter 7 was conducted at The Intelligent Robotics Laboratory at Osaka University as part of Japanese Society for the Promotion of Science (JSPS) Fellowship.

All experimental paradigms were specified by this author, and experimental designs were then developed under the guidance of Prof. Tony J. Prescott, Dr. Abigail Millings and Dr. Ben Mitchinson. Additional support and guidance for the experiment conducted in Japan was provided by Prof. Hiroshi Ishiguro, Prof. Yoichiro Yoshikawa and Dr. Takamasa Iio at Osaka University.

Chapter 2

AAT and RAT: Animal and Robot Assisted Therapies

2.1 Summary

The purpose of this chapter is to establish the theoretical domain that underlies the experimental work described in this thesis, namely the influence that Animal-Assisted Therapy (AAT) has had on the expanding Robot-Assisted Therapy (RAT) field.

To that end the structure of this chapter is as follows. A brief review of the AAT literature is given. The impact of animals on therapy is discussed, and used to introduce the use and impact of robots in therapy. This is done with a particular emphasis on social biomimetic robots in that area. The design of therapeutic robots is then introduced. In this section the focus is on the biomimetic mechanisms borrowed from AAT by RAT. The question of how these mechanisms could be used to tailor robots to RAT users is then raised. The section on design concludes with a brief discussion of how RAT performance is measured.

In an absence of clarity on mechanism in either the AAT or RAT domains the chapter calls for more controlled studies and guidelines that explore what it is specifically about robots that aids positive changes in RAT users.

2.2 Animal Assisted Interventions

A small pet animal is often an excellent companion for the sick, for long chronic cases especially. A pet bird in a cage is sometimes the only pleasure of an invalid confined for years to the same room.

Florence Nightingale (Nightingale, 1969, p. 103)

Animals contribute to their owner's sense of self and well-being, in their need to be cared for and in their ability to display non-judgemental, accepting and attentive behaviour in return for their owner's attention (Burke, 1992). Anecdotally, at least, animals are considered capable of providing their owners with a "boundless measure of acceptance, adoration, attention, forgiveness, and unconditional love" (Bustad, 1981, p. 116). Animals in general then, but in particular companion animals, have been regarded as contributing to human physical and mental health. As such, animals currently play a role in therapy as co-therapist, facilitating a patient or client's existing treatment programme.

Historically the human-animal bond has been employed by healers, via early cultures' use of animal souls and powers in spiritual healing and shamanism, to animism in classical and medieval times, through to the popularisation of animals as agents of socialisation in seventeenth century Europe (Serpell, 2006). However, it was with the introduction of animal-assisted institutional care in the eighteen century that the concept of animals as clinical therapeutic agents was formed.

The Retreat at York, in England, is often cited as the location of the first instance of applied animal-assisted therapy.[1] Founded in 1792 this Quaker run retreat kept small animals, such as rabbits and poultry, in its courtyards for patients to care for (Bustad, 1981, p. 117). However, despite the success of subsequent nineteenth century applications of animal-assisted interventions (e.g., animal husbandry (Bustad, 1981, p. 118)) animals had been largely eliminated from medical settings by the beginning of the twentieth century (Allderidge, 1991).

Since the 1970s there has been a resurgence of animal-assisted therapeutic interventions. This change in practice from animal use rooted in metaphysical ideas about human-animal bonds, towards more prosaic explanations of animals' therapeutic benefits, is almost entirely due to the findings of

---

[1] It should be noted that the term *Animal-Assisted Therapy (AAT)* is applied to a variety of programmes that do not otherwise qualify as therapy in the medical sense of the word. 'Therapy' is defined as the medical treatment of disease, a process that results in curative psychiatric or medical treatment. Some applications of AAT are arguably merely recreational activities, which are defined as pleasant pastimes undertaken for amusement. AAT, not least because its medical mechanistic routes remain unquantified, can be more technically defined as an experience that may provide transient relief or pleasure, but which can otherwise not be described as seriously involved in the course of a human condition (as for example chemotherapy can be). However, for the purposes of this thesis the terms 'AAT', 'RAT', 'animal-assisted therapy' and 'robot-assisted therapy' will be used. This is not to imply that this thesis is advocating that AAT and RAT are unquestionably therapies in the most medical and/or scientific sense of the word, but rather that there use is employed across the literature, and as convenient terms they serve the purpose of this thesis well: in that they can be used to describe the said variety of programmes engaged with alongside traditional treatment programmes which are discussed in the thesis (Kruger et al., 2006). Note that other terms often encountered alongside or in place of AAT are *animal-assisted intervention (AAI)*, or *animal-assisted activity (AAA)*. AAI is often considered an umbrella term encompassing AAT, AAA, and *animal-assisted education* programmes.

a study published in 1980. Friedmann and colleagues studied the life expectancies of cardiac care unit outpatients between 1975 and 1977. In a sample of 92 patients, those who were pet owners had a significantly greater chance of survival at one year than non-pet owners (Friedmann et al., 1980).

Extensive medical literature exists confirming a strong link between social support and improved human health and survival (Uchino et al., 1996). The precise mechanisms underlying these effects are subject to debate, as this thesis will report. However, the benefits derived from supportive social relationships, even those provided by animals, are undoubtable. Any relationship in which a person feels cared for, loved or esteemed is deemed to result in better health and well-being for the individual (Serpell, 2006). This notion is central to the thesis.

Animals are used to complement existing treatments, and one form this takes is in being a social focal point resulting in relaxing effects for the human in the short-term, and improved health benefits in the long-term (Serpell, 2006). Animals can be used to help build rapport between the client and the therapist, for example by addressing the initial nervousness felt by a client in a treatment scenario with their enthusiastic greetings. Therapy animals can also act as catalysts for emotion. Fine et al. (2011) report that for many clients the mere presence of an animal at a therapy session can stir emotions otherwise difficult to express. Whilst there is even benefit to be found in using animals vicariously via role modelling in a therapy session. For example, the relationship between a therapist and their animal may be used to demonstrate to a client the caring nature of the therapist (Fine, 2006).

A typical AAT session with a dog is described by Toufexis et al. (1987), who writes that teenagers attending counselling sessions are more likely to open up if their therapist brings a dog along. A typical session will begin by introducing the dog to the client to aid trust between the client and the therapist. Sometimes the dog will become a conduit of the client's feelings as they are expressed to the therapist. A client might say, "Your dog looks pretty sad," meaning, "I'm pretty sad" (Toufexis et al., 1987, p. 74). Animals can also be used to symbolise the challenges facing a client, for example by discussing a gerbil running on its wheel in terms of a human stuck in a circular thinking pattern, such as that characteristic of PTSD (Davis et al., 1996).

2.2.1 The impact of AAT

Some of the literature reporting the positive impact of AAT relies on anecdotal evidence and there exists limited research aimed at identifying the underlying mechanisms of AAT that produce therapeutic change (Fine, 2006). Empirical work establishing an evidence base for AAT is typically conducted by exploring observable changes before and after an AAT intervention. Thus a typical AAT experimental design may resemble the following: A sample of ten patients attending an Alzheimer's Day Care Centre participated in a repeated measures study comparing three weeks' control activity with plush toy dogs, and three weeks' AAT. Significant decreases in anxiety and sadness scores, and increases in positive emotions and motor activity (measured via psychological metrics and behavioural observations) were found after AAT versus the control (Mossello et al., 2011).

There are a large number of controlled outcome studies in AAT, similar to that by Mossello et al. (2011), which explore the effects of the therapy against a control group on a variety of different complaints. From depressive and anxious symptoms, to non-pharmacological pain management. For instance, AAT when compared against a no-intervention group has been found to lead to reductions in self-reported pain levels in children (Braun et al., 2009). Despite this abundance of empirical work there remains little evidence explaining the mechanisms that facilitate the changes. Kazdin and Fine (2010) report three reviews that convey the status of AAT empirical work. These reviews, detailed below, are indicative of the paucity of AAT studies that not only evaluate controlled outcomes, but further establish the effectiveness of the AAT intervention itself by exploring the precise mechanisms that produce therapeutic change.

In Nimer and Lundahl (2007), 250 studies were reviewed, 49 of which met the inclusion criteria and were subsequently submitted for meta-analysis. Overall, AAT was associated with moderate effect sizes in improving outcomes in four areas: ASD, medical difficulties such as issues with heart rate, emotional well-being measured via such as anxiety and depression scores, and behavioural problems such as aggression. Approximately half of the 49 studies included in the meta-analysis employed a control or comparison group. However, when compared against studies that did not have a control no significant differences were found, leading the authors to conclude that controlled and uncontrolled studies can be legitimately compared against one another to produce a consistent pattern of the positive effectiveness of AAT interventions on a range of conditions and populations.

Souter and Miller (2007) present a review and meta-analysis focussed exclusively on animal-assisted activities (AAA) as a treatment for depression. All studies included had a comparison or control group. Only five studies were identified for analysis. Of these five, one (Panzer-Koplow, 2000) had also been included in the aforementioned Nimer and Lundahl (2007) meta-analysis. Aggregate effect sizes for all studies was medium magnitude and statistically significant. The authors concluded that AAA/AAT interventions are associated with fewer depressive symptoms, as measured via self-reports. However, the authors highlight that only a small proportion of AAA/AAT research meets minimal standards of research design. There is little focus on important physiological measures, such as blood pressure. Whilst once again the studies reviewed are of a before and after comparative design.

Finally, O'Haire (2013) reviewed animal-assisted interventions (AAI) as a treatment practice for autism spectrum disorder (ASD). None of the fourteen studies included in this literature review were included in the Nimer and Lundahl (2007) or Souter and Miller (2007) paper's discussed above. Again, despite unanimously positive outcomes from AAI on multiple areas of functioning known to be impaired in ASD (e.g., social interaction), O'Haire reports many methodological weaknesses amongst the included studies, such as small sample sizes, poorly characterised samples, and limited modes of assessment. Whilst once again a major limitation is the lack of appropriate control conditions. A comparison of an animal against a plush toy goes someway towards studying the impact of AAT on a client, but the question of what is actually being measured in the difference, i.e. what it is about the animal that makes it different to the plush toy, is never asked. Nearly one-third of studies included in O'Haire's review did not have a control, instead implementing simple pre-post comparative measure designs.

Many other reviews exist, but they tell a similar story. AAT appears effective, though the mechanisms underlying that effectiveness are little understood. Although controlled studies of AAT are increasing they are fragmented amongst different applications, resulting in diffuse conclusions which pair the effects of different treatments against different conditions. However, given the nature of AAT as a complementary therapy to be conducted alongside a traditional treatment tailored to the individual it is not surprising to find such a wide range of different applications of AAT being assessed alongside one another, as in Nimer and Lundahl (2007).

This method of analysis results in AAT outcomes from across the board being compared against each other without regard to the variety of factors involved, for example, in the treatments of distinct conditions.

With the emergence of the new field of RAT whose intention is to borrow from AAT methodology in order to provide therapists with more tools to treat individuals who could not otherwise benefit from AAT due to circumstance, it seems more important than ever to build a systematic approach to mechanism assessment.

2.3 The Use of Robots in Therapy

Animals are not always readily available for patients. In many care facilities, AAT is an infrequent scheduled event, taking place once or twice a week for a few hours or less (Stiehl et al., 2005). Some care homes entirely restrict animal visitations due to concerns about disease, allergies, aggressive outbursts from patients or animals that could result in injury, or other discretional reasons. Guidelines exist to minimise transmission of zoonotic pathogens and cross-transmission of human pathogens in clinical settings that otherwise wish to include animals to assist in patient therapy (e.g., Murthy et al. (2015)), but aside from such measures RAT has also been proposed as an easy-care alternative to AAT (Bharatharaj et al., 2015).

The objective of RAT is to create robots with the capacity to act as animal surrogates for individuals who do not have access to animals (Stiehl et al., 2005). The idea is not to replace animals, but to create opportunities for individuals to benefit from the positive effects of AAT in situations where AAT is not otherwise possible.

The use of robots in therapy has some precedent aside from that grounded in the AAT model. For example, the development of autonomous, mobile robots as therapeutic tools for children with ASD is conceptually related to constructionist learning approaches, which focus on active exploration of the environment (Dautenhahn and Werry, 2004). In 1976, Weir and Emanuel published research which employed the programming language LOGO to create learning environments designed to actuate communication in young children with ASD via interactions with a small mobile robot turtle (Emanuel and Weir, 1976). ASD therapies continue to benefit from the use of interactive robots, which help to elicit behaviours from patients, teach skills, provide feedback on performance, and improve patient engagement (e.g., Diehl et al. (2012); Scassellati et al. (2012)). Many of these ASD therapeutic robots are humanoid, such as the KASPAR social robot designed by the University of Hertfordshire's Adaptive Systems Research Group.[2] KASPAR (Figure 2.1) is a robot 'toy' designed to help facilitate social interaction with children at the middle- to lower-end of the autistic spectrum who have inhibited ability to communicate.

---

[2] Note that humanoid robots are considered biomimetic in that they employ understanding from biological systems to solve issues in robotics.

Figure 2.1. KASPAR is a minimally-expressive humanoid robot designed as a therapeutic 'toy' for children with ASD. It was designed by the University of Hertfordshire's Adaptive Systems Research Group.

See also Belpaeme and colleagues, who report on the beneficial effects of applying humanoid social robots in clinical settings. These robots help the development of social bonds amongst children diagnosed with conditions such as diabetes. Robots in these scenarios are used to help support the children through what can be a difficult time in a hospital environment (Belpaeme et al., 2012).

However, the most widespread application of RAT has been in the area of elderly care, in particular dementia treatment programmes that employ animal-like robots. PARO is one of the most active commercial examples of a therapy robot intended to be used in a manner akin to AAT. PARO is an interactive robot, modelled after a baby Canadian harp seal, developed by the Intelligent Systems Research Institute (ISRI), a leading

Japanese automative company. It is marketed as a therapeutic tool for use in nursing home settings. It is sold on the premise that it will "interact with human beings (...) to make them feel emotional attachment to the robot" (Shibata, 2016). It does this by engaging its user with basic capabilities that mimic live animal behaviours: sensing touch, recognising a limited amount of speech, expressing small utterances and moving its head, flippers and tail. The relationship that develops between a user and the robot is built upon the limited reactions the robot makes to the user's spoken and physical actions (Kidd et al., 2006). PARO is designed for, amongst other things, use in therapy sessions attended by individuals suffering from dementia and other conditions of cognitive decline. In such individuals emotional capability does not decline in a one-to-one fashion with cognition (Magai et al., 1996) allowing for meaningful application of psychological and emotional therapy. PARO does not locomote, and is designed to be held and fussed over.

2.3.1 The impact of RAT

If robots are to be used in therapy, it is important to understand the positive impacts they have. As explained above there is a large body of research that exists outlining the positive benefits animal therapy has across a broad spectrum of conditions. AAT is known to lower stress (Allen et al., 1991), reduce heart rates (Ballarini, 2003), elevate mood, and contribute to social facilitation (Collis and McNicholas, 1998). With RAT being proposed as a viable alternative to AAT, one way of understanding the positive impacts of RAT is by comparing the RAT literature against the AAT literature and asking, do robots perform in RAT as animals perform in AAT?

There are plenty of review papers that explore the impact of RAT and AAT on particular conditions (e.g., Filan and Llewellyn-Jones (2006); O'Haire (2013); Souter and Miller (2007)). There are very few papers however that directly compare a RAT and an AAT condition in the same study on the same population. This subsection will explore this literature, and attempt to provide an answer to the question of whether robots perform in RAT as beneficially as animals do in AAT.

According to Filan and Llewellyn-Jones (2006) AAT is gaining popularity in residential care facility treatment programmes. Their 2006 review paper concluded that AAT may ameliorate behavioural and psychological symptoms of dementia. The paper indicates that this could be attributed to an animal's ability to aid with lowering blood pressure and increasing the presence of neurochemicals associated with relaxation and bonding. In five of the reviewed studies the presence of a dog significantly reduced aggression and agitation, whilst others noted significant increases in patient's pro-social behaviours.

The same review looked at preliminary evidence for RAT to achieve similar outcomes (two studies, one focussed on the AIBO (Tamura et al., 2004), a second on the NeCoRo cat-like robot made by Omron Corporation (Libin and Cohen-Mansfield, 2004)). Both studies reported patients responding favourably to the robots, though both samples of residents were reluctant to touch the robots (Figure 2.2). Thus, in this early review the benefits of AAT far outweigh those of RAT.

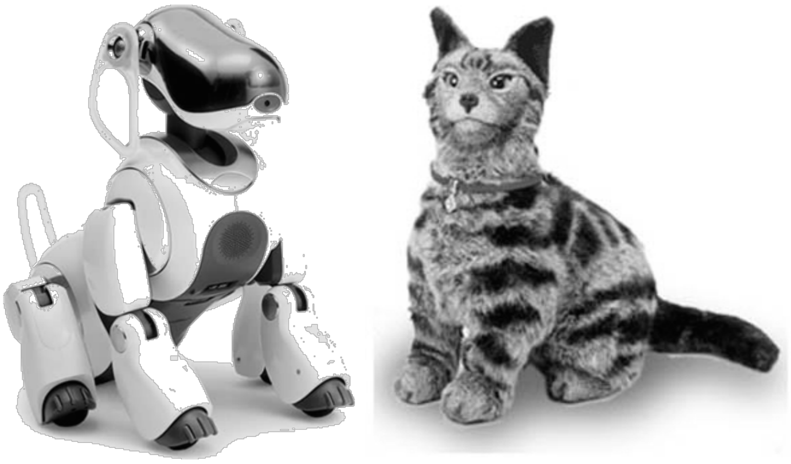


Figure 2.2. Left panel: AIBO by SONY, model shown *ERS 7*. Right panel: NeCoRo the cat-like robot produced by Omron.

In a more extensive 2013 review of RAT as part of dementia treatment programmes (including nine journal publications and twelve conference proceedings), RAT was deemed a potentially useful therapeutic intervention (Mordoch et al., 2013). As with AAT there is evidence that RAT stimulates engaging interaction with other people, producing calming effects, and motivating activity amongst the patients. However, also as with AAT, the review reports that studies are not robust; they have small samples, lack controls, and are often difficult to replicate (see also, Broekens et al. (2009)).

This concern regarding major weaknesses in the design and reporting of RAT studies is echoed in a 2016 comprehensive review paper of robots used in ASD interventions (Begum et al., 2016). Begum, Serna and Yanco note that in order for robots to be seriously considered for clinical settings, they must be considered as an evidence-based practice (EBP). This requires the production of standard research design at a level acceptable by the clinical community, such that systematic evaluation of the evidence from a study can be conducted in order to determine an intervention's effectiveness. This requires, for example, studies which clearly outline how the intervention has been designed, how it can be precisely replicated, and what it is exactly about the robot that can be related to the intervention outcome.

Of particular note are those papers that deal with the PARO therapeutic robot. This robot is the most commercially successful therapy robot in the world, and is used across Europe and Japan in established dementia therapy programmes (Klein et al., 2013). Much of the empirical work with the robot is from Japan, mainly published in IEEE conference proceedings (Mordoch et al., 2013). Whilst some of the studies collect physiology data such as EEG (Wada et al., 2005a) and facial measures (Wada et al., 2005b), comparing pre- and post-interaction figures, several significant issues exist within the literature. It has been reported that the design of these PARO studies are often difficult to ascertain, it can be challenging to decipher which studies are primary and which are from the same study as the focus of the paper, and predominantly the authors of these papers work with the inventor of PARO, creating a conflict of interest that should be taken into consideration (Mordoch et al., 2013).

However, research into RAT is progressing both in number of studies and sophistication of study design. Whilst recognition of the viability of robotic animal substitutes is also growing as commercially available robots such as PARO become more ubiquitous.

It remains that there are very few studies which actively compare RAT against AAT. The first paper to conduct a randomised controlled trial with PARO in a residential care home setting was published in 2013. Robinson et al. (2013) designed a study to explore how the psychosocial effects of PARO could be compared with a live animal control group. It also evaluated the impact PARO had on the social environment of residents interacting with PARO or a Jack Russell terrier. Twenty participants took part in each group, with no significant differences in level of cognitive impairments between groups. Baseline loneliness, depression and quality of life (QoL) scores were taken. Follow-up measures were administered after twelve weeks. After adjusting for baseline scores no main effects of group on changes in self-rated QoL or depressions scores were found. However, loneliness decreased significantly in the PARO group. Residents talked to one another more in PARO sessions compared with normal activities and when the resident dog was present. This finding extends work conducted by Banks et al. (2008) with the robotic dog AIBO. Over a period of eight weeks elderly residents of a care home reported decreases in loneliness after receiving 30-minute weekly visits from either AIBO or a living dog, in comparison with a control group. That robots can be as impactful as animals in reducing loneliness in care homes is an important result. Improvements in loneliness can indicate improvements in other areas of life, as social isolation is related to depression, QoL, and mortality (Gulrez et al., 2016; Holmén et al., 1999; Penninx et al., 1997).

There is an issue of accessibility too. The positive effects of animals in care homes can sometimes be mediated by the decision the animal itself makes as to who it visits. Whereas a robot can be placed with any resident and will consistently respond to them. This highlights the issues of comparing the results of different populations. With care home residents interactions that are more structured and consistent can lead to improvements in the resident's psychosocial behaviours. However, compare this to behavioural analysis of human-robot versus human-animal interactions with healthy populations. Here, the agent being interacted with (robot or dog) has been shown to have a significant effect on tactile engagement from the human (Kerepesi et al., 2006). Healthy humans (adults and children) do not touch robots as frequently or with as much complexity as they do dogs. Yet tactile stimulation is also associated with positive psychosocial outcomes (Kutner et al., 2008). Thus the environment in which a robot or animal is being interacted with, plus the condition that the human is in, contribute to the significance of the outcome measures. Such is the situation in RAT and AAT research. To find out if RAT is as effective as AAT, or if robots perform as animals do in any given therapy, the environment and condition must first be established to ensure valid comparison of studies. At present there is not enough literature to do this (Bharatharaj et al., 2015).

Overall, and in a similar manner to AAT (Fine, 2015), the positive impact of robotic interventions can in part be attributed to physiological effects via the reduction of stress hormones (Suga et al., 2002; Wada and Shibata, 2006), and via improvements in brain functioning. Whilst the overall positive psychological effect of helping to develop social relationships is also fundamental. However, it remains that the underlying mechanisms in both RAT and AAT that are contributing to these favourable outcomes is still largely unknown.

The purpose of RAT is not to provide caregivers with robots that can replace animals in AAT, but to offer an alternative therapeutic intervention for individuals for whom AAT is not otherwise an option, or for whom RAT might offer a better outcome (for example, because a robot is not alive, and thus cannot wander away from a patient). Therefore, in looking for the mechanisms of effect the differences as well as the similarities between animals and robots should be noted. The reasons that a RAT or an AAT intervention is successful, even when being used to treat the same condition in the same environment, may not be equivalent. Different robots and different animals are used in many different kinds of activities, by individuals with different conditions who are living in a variety of environments. Further, because of AAT's client-specific nature there is always the issue of individual differences to take into consideration when reviewing treatment outcomes, and considering those outcomes in comparison to RAT. All of this culminates in an absence of clarity on the mechanisms causing the impact of the interventions.

One issue then is that robots for RAT are being marketed as already equipped to be surrogate animals without the controlled studies to support that claim. This is potentially confusing for practitioners keen to know how to most effectively employ robot tools. It is also unhelpful for robot designers because there is a lack of understanding as to the specific aspects of robots that should be focussed on in order to produce the best possible devices. In sum, the answer to the question of whether robots perform in RAT as animals perform in AAT, is that it remains unknown.

2.4 The Design of Therapeutic Robots

If robots are to produce positive outcomes during therapeutic interventions that are equivalent to those seen with AAT it is necessary to know what forms such robots are required to take. In this section this issue shall be addressed via the following questions.

How should successful therapy robots behave? To answer this it is necessary to know what the mechanisms of AAT are. It is also necessary to know whether one type of robot that behaves in one specific way is appropriate for all situations or not, and if not why not?

How is a RAT robot's performance to be measured? Can the tools used to measure performance in AAT be used equivalently in RAT? Are there general psychology metrics that are applicable to measuring either intervention that will allow researchers to explore aspects of function that can aid improvements in robot design?

2.4.1 Exploring mechanisms of effect in AAT to find biomimetic mechanisms for RAT

A key open question, to the end of understanding the mechanisms of AAT for the betterment of RAT robot design, is how does the nature of the animal effect change in its user? The argument to use human-animal interaction as a source of inspiration for designing social biomimetic robots has been made before, for example in a review article by Miklósi and Gácsi (2012), where the functionality of embodiment and behaviour are cited as mechanisms it would be prudent to aim to replicate in biomimetic robot design. However, extracting such information is not as straightforward as observing an animal and attempting to reproduce an animal in mechanical form. Robot morphology and behaviour should ideally be functional, and in order for function to follow such, the mechanisms of effect must be understood.

For example, in Banks et al. (2008) the authors highlight that the exact mechanisms by which AAT results in decreased loneliness are unclear. The authors continue by hypothesising that attachment may be a mechanism involved. That the emotional bond that supports a sense of well-being, security and closeness in turn results in the physiological changes that drive positive AAT outcomes. In this thesis the role that attachment plays in the development of a relationship will be addressed (see Chapter 3). In this subsection, however, the broader picture of possible mechanisms in AAT, including attachment theory, will be more generally discussed with the emphasis placed on how this understanding can benefit RAT.

General mechanisms of therapeutic action have been suggested in the AAT literature. These fall roughly into two categories. One is the notion that animals have particular attributes which facilitate and contribute to therapy. The other is the idea that it is the process of developing a working relationship with an animal that leads to positive changes in behaviour and cognition via physical interactions which may lead to the development of new skills, or which instills in the individual a sense of personal agency and responsibility. This latter mechanism is seen in AAT that employs farm animals, where the process of managing and caring for animals increases self-efficacy leading to reductions in, for example, depressive and anxious symptoms (Berget and Braastad, 2011).

Of these two mechanisms it is the former that should be the focus of current state-of-the-art (SoA) robotic's design. The acquisition of agency and sense of responsibility described in the second of these proposed mechanisms, which comes from tending to a living animal cannot presently be replicated with a robot.[3] Rather, it is those mechanisms born of the supposedly intrinsic attributes of animals which facilitate therapeutic benefits, that have the potential to be replicated with biomimetic robot-assisted therapy.

In 2006, Kruger and Serpell listed three possible mechanisms via which the mere presence of an animal, with their particular set of attributes, their natural behaviours and companionship, may contribute to the therapeutic process (Kruger et al., 2006). These are, 1. Reduction of anxiety and arousal, 2. Social mediation, and, 3. Attachment theory, transitional objects, and social needs. These shall here be discussed in turn.

Reduction of anxiety and arousal has been attributed to the idea that the mere presence of an animal produces calming effects in humans. One explanation as to why this may be the case is found in the biophilia hypothesis, proposed by Wilson (1984). The hypothesis argues that humans are innately drawn to lifelike and living processes. Evolutionarily speaking humans can increase their chances of survival by focusing on environmental cues, which results in humans being attracted to other living organisms.

---

[3] Some might argue that the tamagotchi phenomenon (e.g., Chen et al. (2005)) could parallel this notion. That the tending to an electronic 'life' could produce similar feelings of caregiving in a user. However, it could also be argued that the psychological outcomes that occur when tending to a device are far removed from those that occur when tending to a living being whose life is dependent upon its caregiver. In farm AAT (also called care farming, social farming or green care farming; the use of commercial farms and agricultural landscapes to promote human health) it is the process of manual labour – feeding, milking, cleaning, etc. – and the subsequent relationship that develops between the individual and the animal that is central to the mental and physical therapeutic process.

However, despite an abundance of evidence, and anecdotal references, that suggest the mere presence of animals is enough to cause physiological and psychological changes in a human, there is no evidence to suggest that this is due to a human's innate attraction to animals in and of themselves. Alternatively, Brickel (1985) argues that learning theory explains the anti-anxiolytic benefits of human-animal interactions. Simply put, interacting with an animal is pleasurable, and pleasurable activities are more likely to be repeated in the future. Animals provide therapy sessions with a calming buffer, diverting attention away from anxiety inducing effects of the session and resulting in more favourable outcomes from the therapy itself. For example, psychological therapy can be difficult for a client who may be addressing issues for the first time, or be doing so in a novel environment, and the presence of an animal can contribute to the client's perception of the environment as safe (Parish-Plass, 2008). According to Parish-Plass (2008) the client's perception that the therapist is making their animal co-therapist feel safe can by proxy contribute to the client's perception that the therapist will ensure that they will feel safe as well. However, Brickel (1985) offers no direct explanation as to what mechanisms are at play which lead to animals inducing such feelings of comfort. Once again this is an explanation of why AAT may work, which lacks an explanation of the mechanism of effect despite purporting to be a mechanistic explanation. The presence of an animal may reduce anxiety and arousal in a client, but that does not answer the question of why that is the case. Is the the demeanour of the animal? Its fur, or trusting expression? The sounds an animal might be making? How an animal moves, perhaps resting its head a certain way, or walking at a certain speed? Without this latter mechanistic explanation it is difficult to design biomimetic properties into a robot that could faithfully replicate this scenario.

Tactile behaviours have been linked to reduction of anxiety and arousal (e.g., Kutner et al. (2008)). Perhaps it is the fur of an animal that elicits touches from a human which in turn leads to the anti-anxiolytic effects? If this can be evidenced then fur could be purposely built into biomimetic therapeutic devices. However, there is no evidence for such an effect. Further, a study published by Shiloh et al. (2003) is suggestive of fur not being a primary factor in anxiety and arousal reduction. In the study 58 non-clinical participants were exposed to a stressful situation (the presence of a tarantula spider they were informed they would have to later hold). State-anxiety was measured at baseline and after the stress and experimental manipulations (using the State-Trait Anxiety Inventory). The experimental manipulation consisted of a 2-minute exposure to one of five groups (random assignment): being left alone with a rabbit, a turtle, a toy rabbit, a toy turtle, or a control group that received neither an animal nor a toy. Participants were instructed to hold and pet the experimental agent, or in the case of the control group was asked to wait, whilst the experimenter left the room. After controlling for participants' attitudes to animals (using the Companion Animal Semantic Differential) results revealed that petting an animal, rabbit or turtle, reduced state-anxiety. The effect could not be attributed to petting the animal per se, as the petting behaviour was not matched in the toy conditions. However, the results appear indicative that both a furry and furless animal can produce equally reductive effects on anxiety.

One mechanism which may be responsible for such a result is the discriminatory effects of myelinated sensory nerve fibres, specifically A-beta afferent tactile fibres in the hands. The pleasant properties of touch are hypothesised to be mediated by unmyelinated peripheral nerve fibres – CT afferents (C tactile fibres) – which project to the posterior insular cortex and signal pleasant aspects of touch (McGlone et al., 2012). However, CT afferents are not found in glabrous skin (McGlone et al., 2007). Thus it may be that stroking something, even with the primarily discriminatory glabrous skin found on the palm, triggers a different pathway prompting an associative learning pattern of stimulation with positive affect (McGlone et al., 2012). That both touch on hairy skin, mediated by CT afferents, processed in the limbic-related cortex, and touch on glabrous skin mediated by A-beta afferents, processed in the somatosensory cortex, are perceived as pleasant goes someway towards highlighting the extent to which pleasure pathways in the brain are not yet fully understood. In sum, though both furry and furless animals appear to stimulate reductions in anxiety, the variety of pathways involved in tactile sensation processing may be indicative of the fact that a specific tactile process is not a contributory factor to the role touch plays in anxiety reduction. Rather, it may simply be the very act of touching that is required to prompt a positive effect.

Perhaps it is the accessibility of the animal that is important, as reported by Robinson et al. (2013), in a controlled comparison of a residential dog and a PARO robot. Here the residents were unable to talk to or touch the dog as frequently as the robot because the robot could be placed in the laps of residents, whilst the dog chose with whom it interacted. In this study residents who interacted with the robot had significant decreases in loneliness over the trial period, and significantly touched and spoke to the robot more than the resident dog (Robinson et al., 2013). No causal links between these observations and results were made in the study, but the presence of them is of note, if not simply because of the existent lack of known causal mechanisms and the need to generate hypotheses regarding what precisely is happening in these scenarios.

Social mediation is another general mechanism proposed by Kruger et al. (2006) via which an animal may contribute to the therapeutic process. An animal can act as a focal point becoming a catalyst or mediator for human social interactions, resulting in increased interactions between humans. The animal is thus a conduit for building rapport between client and therapist, or for prompting communication amongst groups. This mechanism has been studied in RAT (e.g., Šabanović et al. (2013)), and therapeutic robots are currently marketed as social facilitation tools. Perhaps such social mediation effects occur because the agent - be it animal or robot - is non-judgemental (Friedmann et al., 2000). In the case of extracting this non-judgemental characteristic and applying it to robots, is the mechanism effective within RAT because robots are animated but benign? For example, it is argued that robots are effective in ASD therapy because they can be used as a buffer to mediate between an individual with ASD and another person. Robots can provide indirect human-to-human contact, helping to expose a person with autism to tactile interactions in a safe and slow manner before direct human-to-human contact is made (Robins et al., 2013).

The final mechanism proposed by Kruger et al. (2006) concerns the bonds formed by humans and animals. They cite **attachment theory, transitional objects and social need** as notions which help clients to achieve therapeutic advantages in AAT. Much of this theory is based upon anecdotal evidence describing the attachments people form with animals. Attachment theory is a psychological model that describes the short- and long-term interpersonal relationships humans have with other humans, and in other literature the short- and long-term dynamics humans have with other animals and objects as well (see Chapter 3 for a full description of attachment). However, Kruger et al. (2006) report that there is yet to be established convincing evidence of a correlation between attachment and the positive therapeutic outcomes of AAT. In support of the theory some evidence has established a link between attachment to objects and well-being. For example Keefer et al. (2012) report that when a close interpersonal relationship is unavailable, or current social interactions are deemed unreliable, an individual may seek an alternative, non-social security source, which may be an object. It is not a stretch to consider animals in the same manner. Further, if such a psychologically supportive relationship is possible with non-living objects then this mechanism may be replicable with a robot. Whilst it is also interesting to note that even if attachment to the item itself is not the driving mechanism behind the establishment of psychological well-being, it can be posited that an individual's history of attachment-related experiences may be a contributive factor to the process, and thus attachment theory does have something to say about the potential for AAT and RAT to produce therapeutic gains (note that this is an idea that will be returned to in Chapter 3).

A transitional object, as defined by Winnicott (1953) is an object that serves a comforting function over a short period of time. In AAT animals are often cited as producing ameliorative effects in the beginning phases of a therapeutic treatment by serving as a comforting agent. This notion can easily been understood as replicable with a robot.

Social relationships are cited as needs, most famously in Maslow's hierarchy of needs (Maslow et al., 1970). This need, translated as a bond formed with an animal permeates AAT literature. Thus animals, in being long-lasting agents of emotional bond (via the role of attachment figure), or in being short-term transitional agents in the absence of a lasting bond (via the role of transitional object), can serve as agents of social need. One way of conceptualising this is as seeing animals as outlets for nurturing behaviour.

Using agents to elicit nurturing behaviour from clients with the goal of producing therapeutic effects is also applied in non-animal based therapies. For example in doll therapy dolls are given to individuals suffering from Alzheimer's disease to ease anxiety and elicit joy by purposing said individuals with the goal of caring for the object.

Imagine a continuum of caregiving with inanimate dolls at one end, and living animals at the other. Robots can be conceptualised as existing in-between, in a liminal state of being neither alive, nor inanimate, but as agents that display lifelike behaviours (in being autonomous and responsive). Thus the social needs mechanism here described in AAT could be conceived of as being applicable to RAT via sophisticated biomimetic robots regardless of the fact that they are not alive.

However, it may be that animals are uniquely equipped for producing the benefits of the social need mechanism to a population wider than those suffering from late stage degenerative cognitive diseases. An animal *is* alive. The relationship formed with an animal is two-way. Care is given, and received. An animal, anecdotally at least, is empathic (though there is much debate as to the true nature of animal empathy in the literature, e.g., Edgar et al. (2012)). An animal can sense and respond to the feelings of an individual. If this is central to the mechanisms that lead to successful AAT, is it feasible to replicate the equivalent process with a biomimetic robot that is not alive?[4]

All of these general mechanisms listed are evidence of AAT's success.

---

[4] The question of what constitutes a robot – a entity living or nonliving – will be addressed in Chapter 3, Section 3.2.1: Treating robots as the other.

Whilst these last mechanisms which rely on the attributes of the animal suggest that biomimetic robot based therapy may work similarly. However, this is only the case if it can be assumed that these successful AAT mechanisms are not reliant on the therapeutic agent being a living entity. There is evidence that dogs and robots are better at stimulating patients than toy animals, but what is the difference between the dog and the robot? Or perhaps a more useful question to ask for finding successful biomimetic mechanisms for RAT is, what aspects of AAT that produce beneficial outcomes are not reliant on the fact that the agent is alive?

Consider the biophilia hypothesis. Humans have evolved to seek out life. Perhaps that is the important mechanism that results in positive benefits in AAT. Humans see life even where this is none. The psychological phenomenon of pareidolia explains how humans perceive familiar patterns, such as faces, where none such exist (Figure 2.3). Whilst the Media Equation argues that humans treat computers and other technology as if they were alive, responding emotionally and psychologically to computers and media with the same regard as if they were another person (Nass and Reeves, 1996).

Further, it is the individual that is central to the outcome of therapy, as it is ultimately the individual's response that will underpin how effective an intervention will be. A person's psychology drives them to bond with others, and dictates how that bond forms. Thus, perhaps another useful question to ask when exploring mechanisms of effect in AAT, in order

Figure 2.3. An excitedly happy aeroplane: an example of the psychological phenomenon pareidolia.

to find biomimetic mechanisms for RAT, is what role does the individual and their personality, or even their cultural background, play in effecting therapeutic change?

Given the notion that the attachment behavioural system can play a role in driving a person to seek comfort from an object, and the anecdotal role attachment theory plays in AAT, perhaps attachment theory, and the attachment style of an individual, has something to teach designers of robots for RAT. If the biomimetic mechanisms employed in RAT are of a type that stimulate some aspect of attachment in a human user, perhaps that will be sufficient for translating some of the successful aspects of AAT to RAT. The experimental work described in Chapter 4 will explore this.

Exploring mechanisms of effect in AAT to find biomimetic mechanisms for RAT is difficult, not least because AAT is applied as an adjunct to another therapy for the treatment of a variety of different conditions. Animal-like, or humanoid, large or small, furry or not, many design aspects borrowed from AAT and applied to RAT appear relevant to the success of the therapeutic intervention. However, the lack of definition in both AAT and RAT have contributed to fields wherein the answers to some of the most basic research questions, such as what precisely occurs in either therapy that produces positive outcomes, remain unknown. This subsection has raised the further question of the extent to which individual differences amongst users impact outcomes. If robot developers know the answers to this open question of how a successful therapy robot should behave, by exploring the mechanisms of AAT, robots could be built that are better suited for specific uses.

2.4.2 Tailoring robots to users

In designing effective therapeutic robots to produce positive outcomes it is necessary to know whether one size fits all. That is, is one type of robot that behaves in one specific way appropriate for all situations or not, and if not why not? In order to tackle this question a parallel can be drawn with how animals are tailored for use in AAT.

Not all animals are appropriate for all conditions, and for all people. For example, the decision to prescribe equine therapy over canine is made at the discretion of the practitioner and what they know about their patient and their personal needs. Rigorous standardised guidelines for selecting, preparing and integrating animals into treatment programmes do not exist (MacNamara et al., 2015). However, a number of models are available to guide mental health professionals who wish to include animals in their services (for example the Equine Behavioural Profile System (Spink, 1993)). These guidelines focus on the client-specific goal oriented nature of AAT, where specific criteria related to the interactions expected of the animal are taken into consideration prior to animal selection (MacNamara et al., 2015).

Animals are of course alive, and this makes tailoring them to a client a very different process to the decision of whether to use robotic intervention or not. Mental health practice is inherently stressful and intimate. An animal must be able to spend time with an individual, be touched, hugged, directed (e.g., in the case of being cared for or ridden), and be capable of dealing with exposure to inconsistent behaviour during a session. The purpose of the animal in the practice must be clarified (e.g., to engage or to motivate behavioural change); the category and approach of the animal intervention must be decided upon (e.g, for observation, as a focal point, or to build skills with (see Table 2.1)); and the capabilities of the animal must be assessed such that it is correctly matched to a client's needs.

Three quotients make up an animal's capability to work in a mental health

Table 2.1. Animal-Assisted Intervention Categories from MacNamara et al. (2015). These categories offer clinicians a theoretical basis upon which they can consider what attributes an animal brings to an intervention.

Implicit	Explicit	Instrumental
Observing or being in the presence of animals.	Directed observation or simple contact with animals.	Directed activities and interventions between animals and clients.
Animals as part of therapeutic milieu.	Animals as passive therapeutic agents.	Animals as active therapeutic agents.
Redirect attention outwards.	Animals serve to enhance rapport, focus attention outward, enhance assessment, and encourage sensory/cognitive processing.	Focus on process of interventions, skills building, and developing/practicing new behaviours.
Animal contributions are primarily in terms of appearance and/or natural behaviours; interventions are often indirect.		
Presence of animals may enhance rapport/trust.		

intervention: its capacity, which refers to the degree to which it engages with its environment; its skills; and its responsiveness. These are defined in MacNamara's Animal Capability Assessment Model (MACAM). The model includes a fourth quotient too: attributes, which are the physical characteristics that can contribute to the animal's intervention category (MacNamara et al., 2015).

The MACAM model allows a clinician to assess an animal's intangible and physical attributes to aid in selecting the most suitable animal to support the achievement of a client's goals. The process of defining an animal's capabilities for a specific intervention is informative for RAT design. Guidelines such as the MACAM model could help designers to approach the question of how a robot should be tailored for use by offering them an understanding of how an animal's characteristic are applied. For example consider the difference between a chicken and a dog's ability to communicate. A chicken has few behavioural options when it comes to interacting with a human, whilst a dog's range of behavioural engagement is wide. In interventions where explicit communication is unnecessary using a dog as co-therapist may not be the right approach. Physical attributes can have an immediate influence on a client, and when selecting an animal for an intervention a client's cultural, ethnic and gender-based perceptions of animals can make up part of the consideration. For example, as MacNamara explains, using a shepherd-breed canine may be appropriate for an intervention designed for an inner-city youth programme, but may be very wrong when working with refugees or immigrants who have lived in camps patrolled by military dogs (MacNamara et al., 2015). For current SoA robotics, understanding the nuance of a working animal's attributes may not be necessary due to the lack of technological sophistication in SoA. However, as the technology advances this information will be required in order to tailor biomimetic mechanisms in sophisticated RAT programmes.

Another thing to consider when tailoring current SoA robots to users is the interplay between the robot and the user, and the impact that has on potential outcomes (see Section 2.4.3). When an animal is tailored to a client in AAT the temperament of the client is taken into consideration too (Parshall, 2003). The same argument could be made for RAT: tailoring a robot to an individual based on an understanding of that individual's particular differences could go a long way towards maximising effective therapeutic outcomes. Indeed understanding psychological individual differences with the aim of tailoring therapeutic robots could also aid in measuring RAT performance. For example, by understanding how individual differences impact treatment outcomes.[5] This human-centric approach is a key theme in this thesis, and will be returned to in Chapter 3.

2.4.3 Measuring RAT performance

In the absence of clarity on mechanism, the measurement of a robot's clinical performance in RAT must be a key element of the design process. Controlled studies comparing animals and robots reveal that robots are effective at reducing loneliness and anxiety, and help with cognitive decline via social facilitation. The performance of experimental conditions in these studies are predominately measured with change scores, obtained via psychological or physiological metrics taken before and after the intervention.

Currently in RAT as in AAT there exists little evidence exploring what it is specifically about robots and animals that aid positive therapeutic changes. Without clear mechanisms it is difficult to measure performance, because it is difficult to ascertain what factors should be the focus of the

---

[5] However, implementing an understanding of individual differences with the aim of improving robot performance and efficiency of robot use will be difficult until clear mechanisms of effect are understood.

performance measure.

As suggested in the previous subsection there are two main theoretical gaps that could inform better measurement of RAT:

1. A focus on clear mechanisms of effect. With RAT a robot's attributes can be easily manipulated in order to ascertain which characteristics are better than others for a given scenario, a design difficult to conduct in an AAT study as it is difficult to turn on or off an animal's intangible or physical attributes.

2. A focus on the user, and measurements of individual differences in order to discover how particular individuals respond to particular robots. In using established metrics from existing human-other interaction fields researchers can understand the impact that an individual can have on a therapeutic outcome aside from the impact of the therapeutic agent itself.[6]

It is important to stress that such a comparative approach should take into consideration the nature of the robot as a unique social agent unlike a living being or an inanimate object. This shall be discussed further in Chapter 3.

---

[6] For example, in this thesis tactile behavioural measures taken via observations, and metrics developed from attachment theory are used. Observational measures and metrics from attachment theory have been used in human-other interaction studies, as will be explained in Chapters 3 and 4.

With respect to this second human-centric approach to better measurement of RAT, attachment theory and its related psychological metrics is a good exemplar of a human-other methodology that can be borrowed from to help inform RAT and measure robot performance. Attachment theory is an influential theory of personality and development focussed on social interactions with significant others. It is a theory of the long-standing close emotional ties known as ‘attachments’, which are thought to be driven by a behavioural system that is common across all mammals, and evolutionarily programmed to keep infants close to their caregivers and safe from harm (Bowlby, 1969). Attachment theory has illuminated research in human-human, -animal, and -object relationships (Collins et al., 2013) because the attachment system in humans influences emotional reactions and affects. For example, the attachment system, developed in childhood, functions as a method of keeping an infant safe and aware of threat, subsequently safe and supportive environments reduce tension and anxiety, resulting in secure attachment. However, when safe environments cannot be established the attachment system becomes less efficient, and may result in insecure attachment (Bowlby, 1969). A safe environment can be established via the acquisition of a secure base, which has led to the theory’s application in areas of research that explore non-human forms of secure base formation (Collins et al., 2013). Further, the application of attachment theory to models of AAT can inform its use in human-robot assessment.

For example, Barker et al. (2015) report that attachment style may be a moderating variable in how effective AAT can be for individuals. Forty children were randomly assigned to an AAT or active control condition (an age appropriate jigsaw puzzle), and pre- and post-condition measures of pain and anxiety were taken. The Attachment Questionnaire for Children (Sharpe et al., 1998) was also administered. A significant post-condition difference was found between groups for anxiety, with lower scores in the AAT group. Whilst children with secure attachment styles, and who are therefore presumed to feel as though there are in a safe and supportive environment (see above), reported lower pain and anxiety at baseline, with large effect sizes for differences in both anxiety and pain. This demonstrates the impact that an individual's psychology can have on the recorded performance outcome of an intervention. In attempting to establish how well a robot is performing its task, the influence of features such as a user's psychology upon performance scores are required in order to obtain the clearest picture of a situation, and learn how to adjust the application of a robot accordingly in order to ensure that all users of a particular robot receive equivalent benefits.

As in Barker et al. (2015), it is feasible to imagine an equivalent experimental design set up around RAT in order to ascertain the impact that attachment styles could have on outcome measures, including the observable intervention behaviours used to assess outcome and performance.

2.5 Conclusion

Methodology from AAT can be used to explore RAT (and other HRI), by considering the similarities and differences of the agents of therapy (robot or animal). In understanding such similarities and differences a researcher would be better able to discern in what areas RAT could be as useful as AAT. With respect to this, Chapter 2 has outlined several open questions.

1. What are the specific mechanisms of effect of an animal or a robot that lead to positive outcomes in therapeutic interventions?

2. How can mechanisms of effect be categorised such that specific mechanisms can be identified?

3. How important to the therapeutic outcome is due to animals being living beings?

4. When measuring a RAT robot's performance, what exactly is being measured?

5. How important or influential are individual differences to outcome performance measures?

6. If individual differences are having an impact on the effectiveness of a robot intervention, how can robots be tailored to their user?

In the next chapter a new framework will be introduced that considers these areas of overlap between animals and robots, by exploring how humans relate to things in the world around them. This culminates in the proposal of a model of human-other bonds, which aims to provide a consistent conceptual framework with which to describe and analyse human-robot relationships. This is done by considering human-robot interactions against the backdrop of human-other relationships. This will provide the thesis with experimental tooling to help HRI analysis within the domain of comparative AAT.

Chapter 3

Modelling Human-Other Relationships

3.1 Summary

The purpose of Chapter 3 is to consider the tooling that can be used to explore theory and methodology in HRI. This is done by comparing human-robot interactions to existing understandings of human-human, -animal, and -object relationships. The chapter introduces a new framework for describing and analysing human-robot relationships based on a comparative model. Thus, Chapter 3 calls for HRI tooling to be founded upon the notion of the human-agent as central to the analysis, as well as a call for the need to establish a theoretical foundation upon which tooling can be judiciously selected, allowing for clear interpretation of results.

The structure of this chapter is as follows. A conceptualisation of human-other relationships will be introduced, with the emphasis placed on individual differences and their involvement in how an individual relates to their environment. The idea of robots as the other will then be introduced. This is the notion that robots are a form of social agent that can be understood in terms of other social agents, but which remain unique in being liminal – neither alive nor inanimate. Existing attempts to apply methods from human-human interaction (HHI) to HRI will then be discussed. A human factors approach has been suggested as an appropriate method of exploring HRI, however there is still a lack of appreciation within the field of HRI as to how best to utilise such HHI methods given the fact that a robot is not another human, and must therefore be considered differently. The subsequent section explores human-other bonds, and proposes that what is known about human-human, -animal and -object relationships can be expanded to help with the classification of human-robot relationships. This results in the proposal of a framework of human-other bonds that will provide the thesis with a theoretical basis for tooling selection in the reported experimental work.

A note on terms. In this chapter the terms ‘bond’, ‘relationship’, ‘interaction’, and ‘attachment’ are used. Each has a distinct meaning. Here the terms ‘relationship’ and ‘interaction’ are considered equal, defined as the act of two or more agents being connected. A ‘bond’ refers to the connection of two things engaged in a relationship or interaction, and is similarly used. Technically, the term ‘attachment’ should be used explicitly in the psychological sense of the word: a class of long-standing emotional ties driven by an evolutionarily programmed behavioural system intended to keep infants close to their caregivers and safe from harm. However, note that within the literature the term ‘attachment’ is sometimes used to refer to a bond that is actually something other than true attachment. In such cases the term bond is more appropriate, despite not always being used in such cases.

‘Attachment theory’ is a further distinct term, referring to a psychological model that describes the dynamics of interpersonal human relationships, and from which the concept of an individual’s ‘attachment style’ is drawn. This concept is important in the thesis as it is used to categorise individuals, and attempts to help explain individual differences between people and their personal interactions with agents in their environments.

3.2 Introduction to Human-Other Relationships

Human beings form relationships with their environments and the agents within them that are varied and complex. These relationships and their development are influenced by many aspects including love, limerence, practicality, culture, and social commitment. Relationships also fulfil social needs which can be defined as a sense of love and belonging (see the seminal paper by Maslow (1943)). This is not a basic physiological need, but a psychological and emotional one. However, the importance of social needs to healthy functioning in humans is evident in clinical populations that have suffered as a result of minimal social input. For example, early social deprivation can lead to dysfunction in multiple brain regions such as the infra limbic prefrontal cortex, and medial temporal structures such as the amygdala, resulting in long-term neurocognitive and behavioural deficits (Chugani et al., 2001); in adulthood social isolation, such as infrequent participation in social activities and feelings of loneliness, are mechanistic in lower levels of physical health (Cornwell and Waite, 2009); whilst happiness in general is related to how many social connections an individual has, in part due to the way in which friendship networks facilitate the spread of general happiness amongst its members (Fowler and Christakis, 2008). In sum, human beings have a built in requirement for social engagement in order to function effectively.

An individual's relationship with one agent may not be mirrored by another's relationship with the same agent. Such differences between individuals in how they behave and process information is referred to as Differential Psychology. When studying groups placed together randomly, or assigned by biological underpinnings (e.g., disease or gender), differences between individuals in their reactions to control or experimental conditions can be overlooked. Differential Psychology, and the study of individual differences, attempts to avert this by seeking dimensional understanding when exploring groups. That is, looking for dimensions shared by individuals but upon which individuals may differ. Continuum models of personality, such as attachment style and caregiving dimensions (see Section 3.3.1 and Chapter 4), are good examples of such dimensional categorisation.

Acknowledging individual differences and the impact they have on experimental outcomes is important within HRI as it can inform the interpretation of robot performance. For instance, a robot can be designed to be an effective RAT agent by understanding what mechanisms make a particular robot effective for a particular application, however, because of the influence of individual differences upon the development and maintenance of relationships and an individual's interactions with their environment in general, there is no guarantee that a single robot will be consistently effective with all populations.

To make informed predictions about how individuals will respond to a particular robot, human-robot relationships could be conceptualised as lying amongst other classes of human-other interactions. Such a conceptualisation would allow a researcher to borrow tooling from comparative, but more established areas of human-other interaction research to explore the specific impact a particular robot could have on a particular aspect of human psychology. In much the same way that a more refined understanding of the mechanisms of effect will help to improve robot design, a more refined understanding of how humans can be expected to interact with robots, based on what is known of individual differences within human psychology, should also lead to improvements in robot design.

3.2.1 Treating robots as the other

In 2007, Kerstin Dautenhahn wrote, “Robots are not people” (Dautenhahn, 2007). She argued that it is a mistake to assume research from the field of HHI is directly applicable to HRI, such that it will result in honest reliable information. Nonetheless social robots are social agents with whom humans interact and have relationships with. Be those relationships functional, as with an individual having a relationship with their car’s onboard computer, or emotional, as with the owners’ of broken AIBO’s seeking the services of Norimatsu-san, the former Sony employee who now devotes himself to the maintenance of Japan’s ‘pet’ AIBO’s since Sony ceased servicing them (VProBacklight, 2016). As such, the form of social agent that a robot can be categorised as is one possibly comparable to other social agents, but which is ultimately unique. Robots are liminal: neither alive nor inanimate, occupying a position on either side of an otherwise clear boundary in being objects, but ones that behave autonomously and are capable of social interaction (Kang, 2011). Anecdotally there are numerous reports of humans displaying behaviour towards their robots akin to what one might witness from an individual towards their pet dog. For example, descriptions of behaviours such as those reported by Fink et al. (2012) in a paper on anthropomorphic language in online robotic's forums: "Angus' [the AIBOs name] b'day is today. He's had a good day had plenty of dancing, talking and just being a super star," (Fink et al., 2012).

Humans are a gregarious species, with an intelligence rooted in sociality (Dunbar and Shultz, 2007). This social motivation underlies a psychological disposition to anthropomorphise the environment. That is, to ascribe human form or attributes to another: animals, plants, objects, etc. Dautenhahn writes that a human might anthropomorphise a coffee machine, or easily pretend a plush toy horse is alive when playing with a child. As noted in Chapter 2, the phenomenon of pareidolia reveals the extent to which humans see living form in their environment. Minimal morphology, such as a line and two circles, can be enough for an individual to attribute human characteristics to even the most formless of objects (as in Figure 3.1).

It is true that humans feel empathy towards, for example, an embodied AI being abused even when witnessing the act via video feed (Rosenthal-

Figure 3.1. Two circles and a line create a face on a boat: minimal morphology is enough to ascribe human form or attributes to an object (pareidolia).

von der Pütten et al., 2013). However, the relationships that are possible with robots exist in a realm aside from such emotional responses evoked by ‘cute’ pareidolia. Robots elicit behaviours from humans that can be reciprocal, as the robot adjusts its own behaviour to accommodate the responses of the user. As discussed in Chapter 2, some RAT relies on the ability of a robot to elicit an emotional response from its user. For example, as reported in Kidd et al. (2006), the relationship that develops between PARO and its user is built upon the limited reactions the robot makes in response to its user’s behaviours.

The consequences for humans in having socially inclined brains are that human cognitive development relies on social input (for example, see the case of Genie, a victim of severe abuse, neglect, and social isolation (Leiber, 1997)). In return it is the very nature of a human being to be social, to use language, both verbal and behavioural, to interact with the environment, and to project the self onto the other, where the other represents all other relational agents. The importance of being social, and forming relationships, to human life can be understood more fully in witnessing its absence. Such as with Genie referenced above, but also in less severe forms. There are individuals inhibited from being full members of society or culture due to some deficit in social capacity, perhaps as the result of developmental restriction (Kaler and Freeman, 1994), or as a symptom of a condition such as ASD (Baron-Cohen, 1997).

Human beings can respond to agents in their environment with deep sentimentality, and arguably some of this emotion is fictive, created by the individual to address the issue of the social need for love and belonging in the absence of true social engagement (Rodogno, 2015). An individual's relationships are inherently subjective; starting from the self, and projecting outwards to the other in a bid to resolve the need to belong. Epley et al. (2007) argue that anthropomorphism serves the purpose of reducing uncertainty and increasing confidence in predictions, by bestowing upon the anthropomorphised other known quantities of human characteristics and motivations. Non-human agents are unknowns. In the absence of social connections, individuals tend towards using anthropomorphism to create human agents from non-human entities in order to satisfy their own need for social connection (Fink et al., 2012).

Given what is known about human social inclination therefore, an individual's ability to form a relationship with a robot, and be emotionally engaged by one, is unsurprising. Reeves and Nass (1996) are often cited in evidence of a human's disposition to behave socially towards digital artefacts. In 1996 they published The Media Equation, in which was outlined a general communication theory that claims human beings react to digital media as if it were another person. Individuals respond to inconsistencies in the perceived social behaviours of computer programmes with confusion and hurt. One explanation for this behaviour cites Grice's maxims for politeness (Grice et al., 1975). Grice's maxims outline four principles: Quality, Quantity, Clarity and Relevance. A violation of the principles has social significance. In the case of a computer failing to enact a principle when being engaged with by a user, for example in Quantity via not presenting the right amount of information to make a conversation useful, the computer, as one side of a social dyad, will appear unengaged in the interaction, resulting in offence on behalf of the human communicative partner.

In attempting to understand the significance of the relationship between a human and a social robotic agent there are more than generally applicable maxims to be considered. The relationships that an individual develops and how an individual functions socially, are rooted in the individual themselves. It is the personality, the individual differences of a human, which guides their relationships with agents in their environment. General maxims play a role, as The Media Equation can attest, but individual nuances also play a significant role in determining how a person responds to, develops a relationship with, and subsequently maintains a relationship with an agent (e.g., Robins et al. (2000, 2002), see also Caspi and Shiner (2006) for a comprehensive overview of accumulated evidence regarding the influence of personality on life outcomes).

Research in HHI reveals that differences in personality are foundational to an individual's interactions and understanding of the environment. Robots are not people, (Dautenhahn, 2007), but what is known about the influential nature of a person's individual differences upon their relationships with the environment is enough to suggest that HRI should incorporate a methodology for analysis which considers a psychological individual difference's framework. Robots are the other in an individual's environment, much as another person might be, or an animal, or an object. In sum, any relationship a person engages in will be effected by the personality of that person. Thus, despite the fact that robots are not people, explorations of human-robot interactions do have something to learn from existing ideas and metrics in HHI.

3.2.2 Existing attempts to apply tooling from HHI to HRI

There have been some attempts by HRI researchers to apply metrics and methodologies from HHI in their assessment of robot performance. For example, in the application of understanding of gaze. Staudte and Crocker

(2008) were the first HRI researchers to apply HHI psychological findings regarding gaze to an HRI experimental design. In human-human interactions referential gaze has been shown to be instrumental in the planning process of utterance production (Meyer et al., 1998). Inspired by this and other work in the area of HHI gaze Staudte and Crocker employed gaze methodologies to demonstrate that humans react to robot speech and gaze in the same way they would to another human.

HHI proxemics, that is the study of how humans use and manipulate distances, has also been applied to HRI research. Walters et al. (2009) used the Human-Human Personal Space Zone metric (Lambert, 2004), based on earlier proxemics work by Hall et al. (1968), to inform their proposed human-robot proxemics framework by way of comparison. Whilst Sardar et al. (2012) used HHI proxemic's metrics provided by Hall (1966) to demonstrate that participants provide less compensatory behaviour when having their personal space invaded by a human compared with a humanoid robot.

Research focussed on the application of individual differences to HRI however is scarce. A 2006 study by Syrdal and colleagues explored the relationship between individual's personality and robot direction preference (Syrdal et al., 2006). Personality was measured via the Big Five Model (Goldberg, 1999), a domain scale that records emotional stability, extraversion, agreeableness, conscientiousness and intellect. No consistent significant results were obtained to show a relationship between the measured personality traits and preferred approach direction, and the authors in particular note that the effect sizes were too small to be significant in the sample size ($n = 42$). However, high extraversion scores were correlated with higher degrees of tolerance for those robot approach directions rated most uncomfortable. The paper concludes that with the exception of the extraversion correlation, personality traits do not appear to influence the level of tolerance that an individual has to the approach behaviour of a robot.

In 2014, Huang and Gillan reported a study on feelings of attachment experienced by members of various robotics tournaments in which LEGO MINDSTORMS robots were used and eventually dismantled by participants (Huang and Gillan, 2014). Although Huang and Gillan report that the emotional responses to the robots given by participants appeared different from those relationships reported on in human-human or human-pet attachment literature, they did demonstrate significantly positive correlations with the overall affection felt for the robots by participants. Further, affection felt was positively correlated with the time and effort given to robot development during the tournaments. Although not directly indicative of the influencing effect of personality upon robot interaction, this study does highlight how useful HHI tooling, in this case the measurement of conceptual feelings of attachment, can be to the design and development of novel HRI metrics aimed at exploring new areas in a young field.

With respect to methodology, however, it should be noted that without established theoretical foundations upon which to base decisions, the highly interdisciplinary field of HRI is often inconsistent with its selection of appropriate tooling. Consequently results may be prejudiced because tooling was selected on the basis of its expected outcome, and not on the basis of whether it was entirely suited to the task or not. A researcher may well select an instrument from, for example, attachment theory to assess a robot-interaction based on what is hoped to be found, and not on what the constraints of the agents involved will theoretically allow for. For instance, attachment theory can be useful for exploring HRI, and as will be discussed in the following section on human-other bonds (Section 3.3), the general concept of attachment and some of its related measures have been extended for use outside the human-human interaction field by researchers exploring human-animal and human-object relationships. In these cases the fact that there is a non-human half of a dyad being discussed is taken into consideration. Similarly, any application of attachment theory for the purpose of better understanding HRI, should consider the fact that robots are not humans, and subsequent results obtained from studies which include elements of attachment theory as part of their design, should be interpreted accordingly. This can be done by, for example, comparing the application of attachment theory in existing literature on human-object, as well as human-human, bonds in order to form a theoretical basis on which to make predictions regarding how attachment theory might best be applied to understand human-robot bonds. Further, this should be done in light of the particular role of any given robot: its purpose and an example use case, as well as the specific experimental set up in which it is being assessed. In sum, tooling use requires theoretical and practical context to be efficient.

Another issue arising from tooling selection without theoretical backing regards experimenter bias. Sometimes questions are asked in an experiment not because they are appropriate, but because they are typical. This may bias a participant to give a certain result which ultimately supports the experimenter's hypotheses, even if the results are not honestly reflective of the HRI. If a researcher wanted to learn more about a user's perception of a robot they might ask, 'Do you consider the robot to be male, female, or neither?' A researcher might be attempting to discover just how free from gender concepts their robot is, but in asking a question about gender a researcher is giving the participant the expectation that the robot must be gendered at all. Similarly a researcher might wish to know how 'good' their robot is, and ask, 'Is the robot good or bad?' Without a context in which to understand the robot, or knowledge of what a robot is purposed for, as is sometimes the case in HRI when a novel robot is exposed to a participant pool, how can a participant answer the question? Further, if the robot's designer is known by the participant, said participant may answer 'yes' for that reason alone.

Theory and context when selecting tooling for HRI is essential to overcome issues such as these. Further, a theoretical foundation that considers the human and their general relationships as central is one way to provide a conceptual framework upon which the description and analysis of HRI can be laid. Given this premise, the rest of Chapter 3 will discuss human bonds with other things. It shall ask what it is about the human that affects those relationships, which could potentially inform HRI researchers about what to expect from human-robot interactions, relationships and bond formation.

3.3 Human-Other Bonds

Robots do not belong in the same class of living things which humans and animals are, and yet they have capacities for interaction and appearing 'alive' far beyond those of inanimate objects. As such, the proposed conceptualisation of HRI to aid in tooling selection draws from three distinct bodies of literature: the bonds humans have with humans (in particular attachment bonds); the bonds humans have with animals; and the bonds humans feel towards objects.

Once again, with respect to exploring commonalities between, for example, human-animal relationships and human-robot relationships, it should be stressed that the proposed framework is not designed to state that, again for example, human-animal relationships are equivalent to human-robot relationships. As remarked on in Section 3.2.1, robots are not humans, and as analogous as HHI and HRI might be, direct mapping of metrics from HHI on to HRI cannot be expected to provide equivalent results. Rather the framework here proposed is designed to provide an understanding of HRI that allows an experimenter to use similar analytical tools as already exist for analysing one category of relationship, to analyse the other. In doing so a researcher can take advantage of the similarities that, for example, human-animal and human-robot relationships have with one another. In short, the comparative approach provides a researcher with a theoretical foundation upon which to make tooling selection, and to aid appropriate analysis.

3.3.1 Human-human bonds

Psychologists have identified a variety of human-human affective bonds, such as pair-bonds formed in adulthood with partners (Hazan and Zeifman, 1999). An important example class of these are the long-standing close emotional ties are known as 'attachments', and have been analysed by attachment theory (Bowlby, 1969). Attachments are thought to be driven by a behavioural system that is common across all mammals, and evolutionarily programmed to keep infants close to their caregivers and safe from harm. In relation to social needs, attachments, and attachment theory, are believed to underpin how individuals across the lifespan engage with others, in, for example, how they seek out social support, caregive in intimate relationships (Collins and Feeney, 2000), and adjust to novel environments (Nelson and Quick, 1991), all of which are important for physical and mental health (Cornwell and Waite, 2009). This section on human-human bonds will focus on attachment theory, and how the theory relates to individual differences in social interactions.

Attachment theory takes a general systems perspective, positing that an agent's survival is dependent upon the agent's ability to maintain specific variables at specific limits (Marvin and Britner, 1999). This is achieved either via the agent's own abilities, or by the agent combining itself with an outside agent that can regulate the first to maintain its limit requirements (Marvin and Britner, 1999). This is achieved via three connected feedback systems, arranged as organised systems of behaviour based on interpersonal boundaries. Together these systems operate to ensure the survival of the infant agent (Millings and Collins, 2016).

These three systems:

1. Keep infants close to caregivers and safe from harm (the attachment system).

2. Equip the developing infant with the skills required for eventual self-regulation and survival (the exploration behavioural system).

3. Keep caregivers focussed on the infant (the caregiving behavioural system).

As described by Marvin and Britner (1999) these first two systems operate in a single feedback loop. Activation of the attachment system, which is triggered by, for example, environmental and intrinsic cues of threat (darkness, pain, etc.), aborts the exploration system. In response the agent will seek a ‘safe haven’ by regaining proximity to a caregiver. Conversely, satisfaction of the attachment system reactivates the exploration system. The agent then uses their caregiver as a ‘secure base’ from which to explore and learn. The attachment and exploration systems are vital to infant survival, but also continue to be relevant throughout life (Bowlby, 1969). For example, when later attachments are formed with partners and peers (Mikulincer, 1986).

Across the population individual differences occur in these systems. Differences in interpersonal experiences are internalised as cognitive models which guide affect, behaviour and cognition in a manner akin to the influence of personality traits (Mikulincer, 1986). These individual differences affect an individuals’ behaviour, not only with respect to their own attachment relationships, but also with respect to their caregiving behaviour (Millings et al., 2012). For example, the attachment system, which functions as a method of keeping an agent safe and aware of threat, entails that safe and supportive environments reduce tension and anxiety, this results in secure attachment for the agent. However, when safe environments cannot be established the attachment system becomes less efficient, and may result in insecure attachment (Bowlby, 1969). A safe environment can be established via the acquisition of a secure base. This thinking has led to attachment theory’s application in areas of research that explore non-human forms of secure base formation (Collins et al., 2013). This is an example of how the general concept of attachment has been effectively applied outside human-human relationship research.

However, aside from the general concept of attachment which is usefully applied to human-non-human relationship research, it should be noted that within psychology the term ‘attachment’ itself has come to have a very specific meaning which builds upon Bowlby’s theory. Such that, complete attachment to an agent constitutes four hallmarks, which have been clearly identified by Hazan and Zeifman (1999):

1. Proximity seeking – the attached individual preferentially seeks out proximity to the attachment figure, and where possible, will choose to spend time with them.

2. Separation distress – the attached individual is distressed at the prospect of, or in the event of, prolonged or permanent separation from their attachment figure.

3. Safe haven – the attached individual turns to the attachment figure for comfort or support in times of stress, for example when threatened.

4. Secure base – the attached individual’s knowledge that their attachment figure is available (if needed) enables them to explore and mas-

ter their environment.

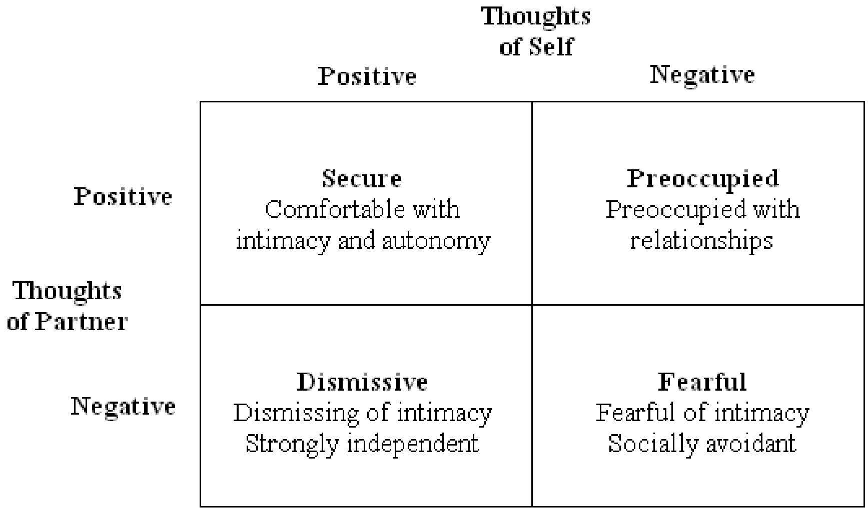
With respect to the application of this specific understanding of attachment being applied outside HHI literature to, for example, HRI, it is difficult to imagine a bond with any existing robot being able to fulfil these criteria. This is perhaps primarily because the state-of-the-art (SoA) in social robotics is not yet up to the challenge of producing a sufficiently convincing agent with which a human might have such an advanced relationship.[1] However, it may be that robots could partially fulfil these criteria. For example, ASD RAT research has shown that interactive robots can mimic interaction games between infants and caretakers (Dautenhahn and Werry, 2004), which play an important role in the development of human social cognition and communication. It is conceivable that such artificially created interpersonal dynamics could elicit some attachment-like responses. As such, Hazan and Zeifman's hallmarks could provide a useful

---

[1] It could be argued that being a 'convincing' agent is not a requirement of an agent towards whom another can become attached. For example, the infant rhesus monkeys in Harlow's maternal deprivation studies (Harlow and Harlow, 1969) were arguably 'attached' to the cloth covered wire 'mothers', with whom they preferentially spent their time with, over wire mesh 'mothers' who sat under a source of warmth (an electric light) and provided the infant with nourishment. However, all of the infants raised by the dummy mothers engaged in unhealthy and abnormal behaviour as adults. Rocking back and forth, clutching themselves, displaying excessive misdirected aggression, and those that had young were further unable to form secure attachments with them, instead ignoring their offspring. Given the maladaptive states within which these experiments were conducted, and the results they produced, there is much to be said as to whether these infant monkeys were in fact truly attached to the dummy mothers, or just suffering from behavioural and neurocognitive deficiencies as a result of social isolation and abuse, which resulted in the observed clinging behaviour from the infants to the cloth 'mothers'. True attachment requires reciprocity on behalf of the agent being attached to, who ensures the infant's survival, and reciprocity does arguably require an advanced agent. This was not the case with the dummy mothers. However, this author would like to acknowledge the complexity of the argument that attachment requires a convincing agent, and the argument's contentious status.

touchstone for social robotics to calibrate itself against even if designing robots to meet all four criteria would currently be problematic. However, this would have to take into consideration what the minimal requirements might be for an attachment figure. Identifying these is another area of research that would be useful to explore when thinking about how humans might form relationships with robots.

A better theory that has been used to understand human-other relationships, which was initially conceived to help explain human-human bonding, is adult attachment style. This is a behavioural repertoire reflective of individual differences in adult attachment behaviour. Described by Bartholomew (1990) an attachment style is manifest in social interaction, and includes how an individual communicates, and how they understand others' behaviour. Bartholomew extended previous work on attachment styles, such as that by Ainsworth and Bell (1970), by proposing a four-category model of attachment based on mental models of the self and others. These models reflect past experiences in interpersonal relationships (including attachment in childhood) and influence present cognitive models of interactions. The four-category model falls along a positive-negative continuum of models of self and other. Individuals with positive self models view themselves as secure and self-sufficient, whilst those with negative self models view themselves as insecure and dependent. Positive models of others are rooted in expectations of others' supportiveness and receptiveness, and result in a belief that relationships and intimacy are worthwhile. Conversely, individuals with negative models of others see relationships as



		Thoughts of Self	
		Positive	**Negative**
Thoughts of Partner	**Positive**	**Secure** Comfortable with intimacy and autonomy	**Preoccupied** Preoccupied with relationships
	Negative	**Dismissive** Dismissing of intimacy Strongly independent	**Fearful** Fearful of intimacy Socially avoidant

Figure 3.2. Batholomew's Four-Category Attachment Style System (Bartholomew, 1990).

unrewarding and have avoidant orientations towards intimacy (Guerrero, 1996). Bartholomew (1990) created an interaction model of these two concepts of self and other to create four quadrants representing four distinct attachment styles (Figure 3.2).

Subsequent work on the model has drawn greater attention to the model of the self as reflective of anxiety and the model of other as reflective of avoidance (e.g., high scores on a negative model of the self correspond with higher anxious feelings, see Bartholomew and Horowitz (1991)).

The four categories can be understood thus:

1. Secure: comfortable with intimacy; confident and self-sufficient; possesses positive expectations about relationships.

2. Preoccupied (also known as Anxious-Ambivalent): craves excessive intimacy; lacks self-confidence; requires relationships to fulfil dependency needs.

3. Dismissive (also known as Dismissive-Avoidant): uncomfortable with intimacy; compulsively self-reliant; does not think relationships are necessary.

4. Fearful (also known as Fearful-Avoidant): fears intimacy; lacks self-confidence; desires closers relationships but fears rejection.

Adult attachment styles, thus conceived as manifest in interactions between a human and another agent, including other non-human agents such as animals and objects, have been used to inform human-other relationship research. For example, in human-animal relationship research, the relationship a client has with an animal co-therapist can be usefully informed with an understanding of the client's attachment style, such that a therapist can then better tailor a particular therapy to a client given the client's predicted behavioural patterns (see Section 3.3.2). Comparatively, it is therefore conceivable that knowledge of a client's attachment style could also help inform a therapist when applying RAT.

In sum, the literature on human-human bonds is extensive and detailed. It provides researchers with an understanding of how humans attach to other humans, and with theories such as that of adult attachment styles to help explain individual differences in relationship formation and maintenance. With respect to the extension of this research to human-robot bond research, although it is arguably not conceivable for an SoA robot to fulfil the four hallmarks of attachment (Hazan and Zeifman, 1999), there is scope for knowledge of attachment theory to inform how a human and another – such as a robot – will engage. The following sections will discuss human-animal and human-object bonds, playing particular attention to how attachment theory has been expanded to help inform each of these areas of research, and how by extension the theory could also be used to help inform human-robot interactions.

3.3.2 Human-animal bonds

In analysing social robots, drawing on the human-animal bond literature is perhaps more helpful than drawing on studies of human-human relationships, since animals and robots share some similarities. For instance, animals, like social robots, are often owned by humans and yet are also more interactive than most possessions. For the purposes of this thesis, which is focussed on social biomimetic, and in particular animal-like, robots, this set of bonds is of most interest.

Human-animal bonds have been examined as attachments, utilising attachment theory to explore human relationships with pets in particular. However, these human-animal relationships tend to be based on different working models to human-human relationships. For example, in Kurdek (2008) the four hallmarks of attachment (Ainsworth and Parkes, 1991; Hazan and Zeifman, 1999) are used to model a new conceptualisation of human-dog attachment-like relationships. In the Kurdek (2008) model, human attachment figures are oriented towards other humans in line with the four hallmarks to a greater extent than pet dogs, however dogs were regarded as equipped to serve as secure bases nonetheless, and further, dogs also exhibited the required features of proximity maintenance.

Other researchers have incorporated attachment theory into the human-animal bond literature via the application of attachment styles to an understanding of human-animal interactions. As described in the section above, attachment styles are reflective of individual differences in attachment behaviour.

An example of the application of the attachment styles, for example those described above by Bartholomew and Horowitz (1991), to human-animal bond literature can be found in Colby and Sherman (2002), who compared how attachment style moderated the impact of a dog visitation programme in 52 elderly residents of an assisted-living facility. Pre- and post-dog visit comparisons of mood were taken, after controlling for the physical status of the participants. The Relationship Questionnaire (Bartholomew and Horowitz, 1991) was administered to ascertain adult attachment style, whilst the Mood Report (Diener and Emmons, 1985) assessed emotional state. Results revealed that secure style related to increases in positive mood after dog visitation, whilst anxious-ambivalent style related to increases in positive mood and decreases in ratings of depression. Further, fearful-avoidant style related to increases in depression after dog-interaction. The Colby and Sherman (2002) study was the first to highlight the relevance of attachment style in understand older population-dog interactions.

Further, Zilcha-Mano et al. (2011) apply an understanding of attachment theory to AAT by considering a client's unmet attachment needs, individual differences in attachment security, coping, and responsiveness to the therapy. Their model is concerned with the goodness of fit between a client's individual differences in attachment style and the development of a relationship with an animal co-therapist by benefiting from an understanding of the nuances of attachment style and the impact these have on an individual's ability to interact with others. Arguably therefore, knowing an individual's attachment style has the potential to inform a clinician of the best ways to foster the development of more adaptive behavioural patterns, thus resulting in better treatment outcomes. Indeed, identifying human attributes such as attachment style which contribute to successful human-other interactions (including -robot) could aid in the development of a behavioural framework that allows full dyadic potential to be realised (for example, as proposed by Payne et al. (2015) for the dog-human dyad).

Other human-animal interaction literature (for example, Walsh (2009)) has discussed 'attachments' to pets, but without using the term 'attachment' in the technical sense defined by attachment theory (i.e. as an evolutionarily-driven emotional tie to a caregiver). Crawford et al. (2006) highlights how researchers have begun to draw upon traditional attachment literature to study human-animal bonds, but they stress that the application of scales for measuring human attachment/bonding to animals may not yield results congruent with attachment theory. However, they stress that despite the differences in the two relationships – human-human attachment and human-animal bonding – examining the human-animal bond via attachment theory can result in useful information. This is helpful when considering the application of attachment theory to HRI. As in the human-animal bond literature an inconsistent use of the term 'attachment' may be confusing, and human-robot relationships would be misconstrued if the term 'attachment' were applied loosely. Here, therefore, the term attachment is reserved for relationships which explicitly feature all of the hallmarks described by Hazan and Zeifman (1999). However, regardless of the semantic issue the use of attachment theory remains a valuable framework within which to explore human-animal relationships as long as the differences between the agents involved in the discussion are kept clear. A human and an animal are different and offer an individual different sorts of interpersonal interaction. Thus, although an understanding of attachment theory, and an individual's attachment style, can help inform how a human-human or a human-animal relationship might unfold, the relationships observed may not always been reflective of each other (Carr and Rockett, 2014). For example, as Zilcha-Mano et al. (2011) note, individuals with highly avoidant human attachment bonds are not always subsequently predisposed to form avoidant attachment bonds with animals (Carr and Rockett, 2014). Understanding this is potentially very informative for understanding how to apply attachment theory to human-robot relationships, an idea which will be explored in Chapter 4.

3.3.3 Human-object bonds

The emotional ties that people have with favourite or sentimental items is another area in which the word ‘attachment’ is often used. As with the human-animal bond literature there is debate as to whether an object can fulfil all the hallmarks of attachment. Thus, within human-object relations literature the term ‘bond’ is a more appropriate characterisation of the dynamic, with attachment theory being applied to help frame the arguments.

In practical terms some researchers of human-object bonds have overcome the issue of whether the object to which a person has an attachment must also have feelings if the relationship is to emulate a human-human attachment style interaction. This is done by re-interpreting classic attachment theory within human-non-human bonding interactions (this argument is had in the human-animal bonding literature too, where the question of whether the animal is ‘attached’ to its owner is also raised (Crawford et al., 2006)). For example, by using adult attachment style to predict levels of attachment feeling towards possessions, as in Kogut and Kogut (2011); or, as with the case of human-animal attachment research discussed above, by proposing models of attachment to objects which feature some – but not all – of the classic hallmarks of human-human attachment (Carr and Rockett, 2014); or, as in Keefer et al. (2012), by exploring ‘attachment’ to objects under the proviso that they offer a secure base *in lieu* of absent social connection. Indeed, the clear hallmarks outlined by Hazan and Zeifman (1999) allows researchers to distinguish between true attachment relationships and other types of bond (such as that with a close possession or an animal).

Other researchers have applied attachment theory to human-object bond research by examining how objects serve as extensions of the self, either as literal tools enabling the owner to do things which would otherwise not be possible, or as symbolic extensions of the self, as when a uniform or trophy allows an individual to adopt an altered sense of self (Belk, 1988), or alternatively to provide coherence in self-narratives by serving as loved objects which allow their owner to symbolically support a self-identity (Ahuvia, 2005). This disposition for an individual’s sense of self to affect their attachment to a personally valued object has also been recorded in children (Diesendruck and Perez, 2015).

It is also worth noting that bonds to objects are often linked to attachment relationships with other people, and can therefore become imbued with attachment-like qualities. Weller et al. (2013) utilise attachment theory to discuss a person’s strength of bond to a mobile phone as a predictor of phone use while driving. Here the phone represents access to attachment figures. Object use in such a scenario falls into a grey area between attachment relationships and object bonding. This highlights the difference between the attachment that an individual has to the people they access via their phone, and the relationship that the individual has with their phone (as the facilitator of their attachment behaviour). Arguably, the object's role in the attachment relationship is integral to the individual's feelings towards that object. This is also seen in behaviours such as the holding on to personal mementos when displaced – facilitating continuity of life despite the loss of a home (Parkin, 1999); or, as discussed in the previous paragraph, the use of inanimate objects to represent an extension of self-personality (Kiesler and Kiesler, 2004).

Studies of human-object bonds are particularly interesting for the study of human-robot relationships because of the insight they offer into what could be termed quasi-attachment – bonding to an object which is associated with an attachment relationship – either through function (as with a phone enabling communication with loved ones) or reminiscence (where the object is a memento of an attachment relationship). It is possible to conceive of social robots as serving such a role as they also feature both aspects: being an owned object, and also allowing a user to connect to loved ones if, for example, they were Internet-enabled.

Considering these applications of attachment theory within the human-object bond literature the aim here is to utilise existing literature to expand a model which describes where along a spectrum of bonds various human-robot interactions may lie. In order to do this consideration of the human-robot bond in relation to those bonds formed between humans and other agents of interaction is useful.

3.4 A Framework for Conceptualising Human-Robot Relationships

Thus drawing from the literature, briefly summarised above, this chapter proposes that human-robot relationships could be analsysed in terms of their similarities to different types of existing bond with other humans, animals, and objects (Figure 3.3; Collins et al. (2013)), providing researchers with, for example, a unified approach to method selection.

Current relationships between humans and social robots perhaps have most in common with human-animal bonds, but commonalities with human-object bonds should not be overlooked, and there is the potential to at least mimic some of the features of human-human bonds.

For any given robot, its exact location within this relational space will depend to a large extent on the context in which it is to be used. In comparison to this framework, recent research in HRI has tended to occupy one of two extremes. At one end is the idea of robots as objects towards which attachment and love should be directed, with the possibility that such human-robot love could be bi-directional (Samani et al., 2011), and at the other sits the view that robots should be seen as mere objects of ownership, things with the potential to facilitate human-human social bonds but not to be bonded with (Wada and Shibata, 2007). However, as robot capabilities advance, they could be developed across the full spectrum outlined here. This framework therefore has the potential to underpin the full range of robotic technologies society could expect to encounter in the coming decades.

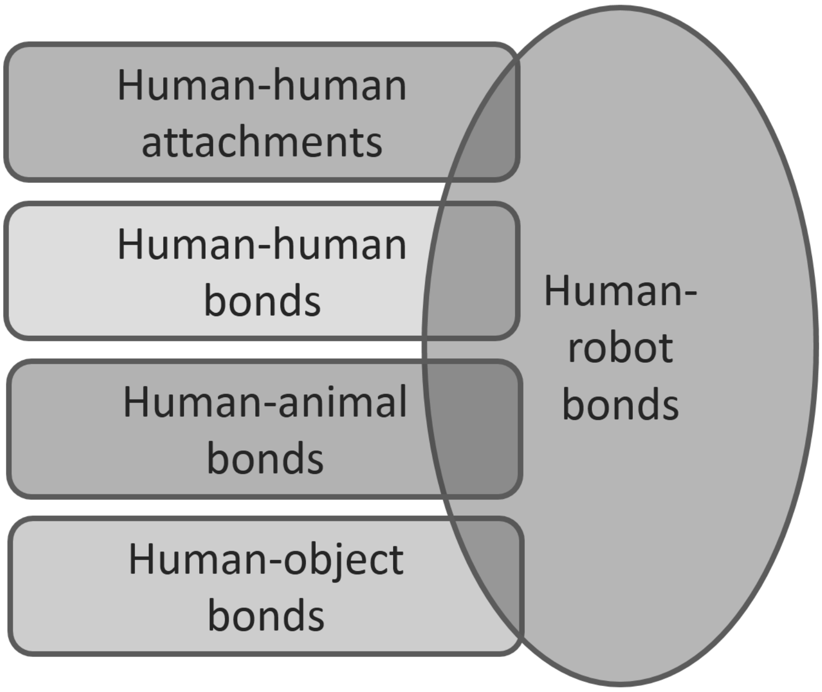


Figure 3.3. Proposed model of human-robot bonds.

3.5 Conclusion

Existing literature on human-other relationships provide established theories and methodologies for exploring how individual differences impact upon a person's interactions with the world. In borrowing from the human-other literature human-robot interactions can be conceptualised as existing somewhere within this relational space. Given that, it is then possible to provide a consistent framework with which to describe and analyse human-robot relationships, which considers the areas of overlap between a particular robot and its living, or static-object, counterpart. Such a clear format lends itself to the production of more precise HRI experiments. Rather than simply appropriating psychological metrics to explore in any particular experiment this framework offers a solid comparative theory to found the selection of metrics upon.

The experimental study reported in Chapter 4 will use this framework to apply attachment theory to the analysis of a human-social biomimetic robot interaction. In AAT there is an open question regarding the extent to which the therapy needs to be tuned to the personality type of the patient, whilst within RAT the question of what parameters of personality are important to the successful outcome of an intervention remains unknown.

Chapter 4

PARO: Felt Security, Physical Engagement and Individual Differences

4.1 Summary

In Chapter 2 the impact of AAT was discussed and used to introduce RAT. The chapter described multiple ways in which AAT has been empirically tested, and highlighted the absence of clarity in mechanisms of effect in both domains. Chapter 3 proposed a comparative framework with which HRI (including RAT) could be explored by drawing on existing metrics applied in human-other interaction studies. The goal of this framework is to facilitate the development of a theoretically grounded methodology that can be employed to answer questions surrounding the impact that human-robot relationships can have.

The work presented in this chapter brings these two ideas together. Previous work in AAT has noted the importance of feelings of care, love and self-esteem to the success of a therapeutic animal intervention. Thus in the empirical work presented in this chapter the effect on healthy participants' felt security (FS), which is a measure of care, love and self-esteem, was taken before and after spending time in the presence of a biomimetic robot designed to be used in RAT (the PARO therapeutic device). This pre/post-condition comparative paradigm was borrowed from AAT, wherein it is used to test the impact of animal interventions. In the event that PARO was found to be having a positive impact on FS, this study also sought to discover *why* PARO might be having a positive impact on its user: that is, a mechanism of effect was also sought.

Drawing on the framework proposed in Chapter 3, metrics were selected to explore how an interaction with PARO might be impacting a user's FS. The first focussed on behavioural observations, a metric frequently employed in both human-human and human-animal interaction studies. Touch is arguably an important mediating effect in AAT, with physical engagement between a human and an animal cited as a mechanism through which positive effects of that engagement are had. Given this, the present study asked whether the level of intimacy with which an individual physically engaged PARO could be correlated against observable changes in FS. The second metric focussed on measuring individual differences in caregiving and attachment styles via questionnaires. These differences have been shown in human-human interaction studies to influence an individual's physical engagement with other social agents, and have been employed in human-animal interaction studies in an attempt to better understand how individuals might respond to animal co-therapists. The present study sought to discover whether correlations between physical engagement with PARO and changes in FS were mediated by an individual's caregiving and attachment styles. Past research suggests that individuals with deactivated or avoidant styles are less physically intimate with social agents than individuals with hyperactivated or anxious styles.

Before the main study was conducted a pilot was run in order to explore the idea that attachment and caregiving styles can be linked to observed physical engagement behaviours, measured via categorisation of touch-types, in an HRI context. The pilot study also served to help develop a coding scheme for measuring physical engagement.

Despite pilot results which appeared to indicate that individual differences in caregiving style were indicative of an individual's physical engagement with PARO, the main study revealed that individual differences in both caregiving and attachment styles were not predictive of the coded physical engagement behaviours. Thus a mediation analysis was not conducted.

However, the present study did show that after controlling for these individual differences, FS did increase after interaction with PARO. Further, how a participant physically engaged with PARO did impact their FS change score. Individuals who spent the majority of their interaction time not touching PARO had lower FS delta scores than individuals who spent the majority of their interaction time touching PARO in a more proximate manner (FS delta is a measure of post-interaction FS score minus pre-interaction FS score). Thus it appears that how a person engages with PARO does impact how that interaction will influence their FS. However, the mechanism driving those differences in touching behaviours and FS delta were not captured by any measures taken in this study.

The chapter concludes by stating a broader implication: If the PARO therapeutic robot can have an equally positive effect across a range of individuals, such RAT agents may be considered generally effective and not restricted to only benefiting certain subgroups. This correlates with AAT literature in which the generic benefits of AAT are promoted as applicable to a range of patients, even when taking individual differences into account.

4.2 Introduction

Chapter 2 discussed the positive impact of AAT on a multitude of conditions, from depressive and anxious symptoms (Nimer and Lundahl, 2007), to non-pharmacological pain management (Braun et al., 2009). Comparatively RAT has been employed as a successful intervention in similar areas (e.g., Banks et al. (2008)). The ability for an interaction with a social agent to positively impact an individual's emotional well-being and sense of self have been cited as possible explanations for the success of AAT and RAT interventions (Banks et al., 2008; Burke, 1992). For example, in their comparative study of the effects of AAT and RAT on loneliness, Banks et al. (2008) state that the negative symptoms of loneliness are possibly alleviated via the development of an emotional bond that supports a sense of closeness, security and well-being, which in turn results in the physiological changes that drive positive AAT outcomes.

Loneliness is associated with symptoms of depression, and positive emotional well-being is mechanistic in ameliorating symptoms of depression, as well as anxiety (as reviewed in Nimer and Lundahl (2007)). The benefits derived from supportive social relationships, including those derived from agents which engage an individual's social self (such as animals, but also potentially social biomimetic robotic agents too) occur because, as stated in Chapter 2, any relationship in which a person feels cared for, loved or esteemed is deemed to result in better health and well-being for the individual (Serpell, 2006).

The present study, which employed the RAT agent PARO (see Section 4.3.1), focussed on exploring whether spending a period of time alone with a social biomimetic robot could increase feelings of care, love and self-esteem in the user. In the event that it was, the study was secondarily focussed on discovering mechanisms for such an effect. This method for exploring the positive impact the RAT agent PARO has upon healthy users is to date unexplored.

Feelings of care, love and self-esteem, as well as safety, are components of felt security (FS) (Bowlby, 1969; Holmes and Rempel, 1989; Mikulincer, 1986; Murray et al., 2000). This can be measured by the Felt Security Scale (FSS; Luke et al. (2012)) which captures the full dimensionality of FS. The FSS was developed as an attachment relationship priming manipulation check for use in human-human interaction (HHI) studies, and has thus far not been used as a measure in an HRI study. However, FS, and related constructs such as self-esteem and safety, are established measures in human-human and to a lesser extent human-animal interaction studies (e.g., felt security in close relationships (Murray et al., 2005) and affecting social interaction for children with ASD via interactions with service dogs which engender feelings of safety (Burrows et al., 2008)). According to the framework proposed in Chapter 3, it follows that as these measures have been used in human-other interaction studies in order to learn something about the relationships therein, they may also be employed similarly in an HRI context in order to learn something about human-robot relationships too. Further, the importance of the construct of FS and its related dimensions of care, safety, love and self-esteem to the success of AAT and RAT interventions (for example, in tackling loneliness and depression) are indicative of the role FS has to play in understanding the positive impact of a human-robot relationship. As such the FSS was selected as the interaction impact check for the present study.

In AAT before and after change score paradigms have proven useful in answering questions about the general effectiveness of an intervention on constructs measured similarly to FS, such as state-anxiety (which measures tension, nervousness, worry and apprehension via self-report questionnaire). For example Shiloh et al. (2003) measured changes in state-anxiety levels via the State-Trait Anxiety Inventory (STAI; Spielberger et al. (1970)) before and after an animal exposure interaction period in order to obtain a change score (in this thesis, referred to as the delta score) and assess the impact of the animal interaction. Thus, a similar paradigm was employed for the present study.

Two further methods with precedent in human-other interaction studies were also selected to explore how an interaction period with PARO might be impacting a user's FS. These are behavioural observations and individual difference's metrics. These shall here be discussed in turn.

4.2.1 Behavioural observations of physical engagement via touch

In order to begin to address the question of what mechanisms might be driving differences in FS delta scores, behavioural observations of participants' physical engagement with PARO were taken.

Touch is arguably an important mediating effect in AAT, with physical engagement between a human and an animal cited as a mechanism through which positive effects of that interaction are had. Barba (1995) reports that though animals need not be touched in AAT sessions for some effects to occur, stroking animals does appear to relax people (see also Miller and Ingram (2000) for the relaxing effects of stroking dogs in AAT). With respect to the cardiovascular system the blood pressure of humans drops whilst stroking dogs, as does the blood pressure of the dog being stroked (Wolff and Frishman, 2004). In a study of 45 residents of a long-term care facility randomised into three groups (no AAT; AAT once/week; AAT three times/week; $n = 15$/group) over a 6-week period AAT significantly reduced loneliness scores in comparison with the no AAT group, where the AAT interaction periods were administered with full engagement protocols: including holding, stroking and grooming the therapy dog (Banks and Banks, 2002).

Whilst such reported positive effects of physical engagement with an animal via touching behaviours are often anecdotal, or minimally explored in a controlled setting, it remains canon within AAT that the act of stroking an animal improves patients' feelings of self-esteem and helps patients confined to clinical settings to feel calmer (Connor and Miller, 2000). Research in social neuroscience on the impact of interpersonal touch and its importance to inter-agent affiliation and social bond formation may offer some explanation to such canon (for example, see Löken and Olausson (2010) for review). CT (C tactile) afferents are a type of mammalian, unmyelinated, low-threshold mechanoreceptors (Zotterman, 1939). Although only found in the hairy skin of humans, and not the glabrous skin of the palm used for stroking, evidence indicates that signalling in such CT fibres via soft, gentle touch, activates somatosensory areas S1 and S2 as well as insular cortex, notably the posterior part of the contralateral insular cortex (Olausson et al., 2010). Whilst the role of such pathways to stroking-acts are little understood, there is growing evidence for the role of CT fibres in oxytocin release (Ellingsen et al., 2014). Oxytocin is the neuropeptide commonly associated with the communication of emotions and strengthening of social bonds. There is much work left to do in this area, but results overall point to mechanisms of touch being key to understanding emotional processing in humans. Given this, the present study asked whether the type and amount of touching behaviours with which an individual physically engaged PARO could be correlated against observable changes in FS. This was achieved via the development of a coding scheme which categorised touching behaviours along a detailed continuum of proximate to distal observed behaviours. This produced a set of scores for each participant that conceptualised how they had physically engaged with PARO during the interaction phase. The development of this coding scheme is discussed in detail for the pilot in Section 4.4.1, and for the main study in Section 4.5.1.

It was predicted that participants who spent the majority of their robot-interaction time physically engaging with PARO in more distal touching behavioural (for example, not touching PARO at all) would have lower pre-interaction FS scores, and smaller FS delta scores overall, than individuals who spent the majority of their interaction time physically engaging with PARO with more proximate touching behaviours (such as stroking PARO).

4.2.2 Individual differences in caregiving and attachment

The other method selected to explore how an interaction with PARO might be impacting a user's FS was founded on the principle that PARO was designed to elicit caregiving behaviours from its user. PARO's ability to basically engage its user, and its static nature (it does not have legs and cannot locomote), require it to be held and entice a user to pet it much like a lap-animal. If physically engaging PARO with more proximate touching behaviours does lead to greater FS delta scores then the question arises as to what might drive an individual to physically engage with PARO in a certain way, and further, might there be differences between individuals and their touching behaviours such that certain populations will consistently derive greater benefits from an interaction with PARO than others? Thus the present study also sought to discover whether correlations between physical engagement with PARO and FS delta were mediated by individual differences in a participant's caregiving and attachment styles.

In HHI literature a propensity to physically engage others with proximate behaviours, such as those deemed intimate or caressive, has been linked to individual differences in caregiving and attachment styles. In Chapter 3 (Section 3.3.1) the attachment system was discussed in relation to a series of other behavioural systems, including the caregiving system. The understanding of attachment styles and their related behavioural repertoires has its foundation in the Strange Situation assessment protocol[1] from whence the three key attachment styles, of secure, insecure avoidant and insecure ambivalent/resistant, come (Ainsworth et al., 1978).

Despite subsequent debate as to what the Strange Situation was in fact measuring (Braungart and Stifter, 1991; Clarke-Stewart, 1989; Crockenberg, 1981; Egeland and Farber, 1984) the research has been extensively used by interpersonal-psychologists to understand agent-agent relations

---

[1] The Strange Situation protocol occurs over a 20-minute interaction in a laboratory setting in which a one-year old infant and its mother are observed. Over the course of the observation a second adult, a stranger, is introduced. The mother then briefly leaves before returning. A second separation then follows in which the infant is left entirely alone. Finally mother and stranger return. In the presence of mother the infants actively explored the playroom. However, individual differences were observed between infant reunion behaviours. These evolved into the Strange Situation classification system of three key attachment styles: secure (the majority, who sought proximity upon reunion), insecure avoidant (avoiding mother upon reunion) and insecure ambivalent/resistant (cried and wanted contact with returning mother but were fussy upon being picked up) (Ainsworth et al., 1978).

(e.g., Bartholomew (1990); Bartholomew and Horowitz (1991); Brennan et al. (1998); Hazan and Shaver (1987); Shaver et al. (2010)). At their core caregiving and attachment styles are explored via their associated behavioural patterns, and the differences therein.

The established two-dimensional understanding of attachment styles is arranged around interacting continuums of anxiety (fear of rejection and abandonment) and avoidance (discomfort with closeness and dependence upon others). An analysis of attachment style measurement literature by Brennan et al. (1998) resulted in the production of the now well established Experiences in Close Relationships (ECR) measure of adult attachment. The ECR measures the two dimensions of avoidance and anxiety, and combines the resulting continuous scales to form four attachment styles: secure (low avoidance and low anxiety), preoccupied (low avoidance and high anxiety), dismissive-avoidant (high avoidance and low anxiety) and fearful-avoidant (high avoidance and high anxiety) (Figure 4.1).[2]

Attachment behaviours are understood as responses to encounters with threat and requirement for protection. Related to that are caregiving behaviours, which are reactions to the distress signals and attachment behaviours of others. Individual differences in caregiving orientations and strategies are measured with the Caregiving System Scale (CSS; Shaver

---

[2] The ECR, and other measures of attachment, should always be considered in dimensional terms. There is no evidence for precise attachment typology (Fraley and Waller, 1998). The space in which the attachment – and caregiving – styles are conceived is two dimensional and individuals fall on that space in continuous terms, as soon as attempts are made to use precise typological terms the dynamism of the continuous scale is lost.

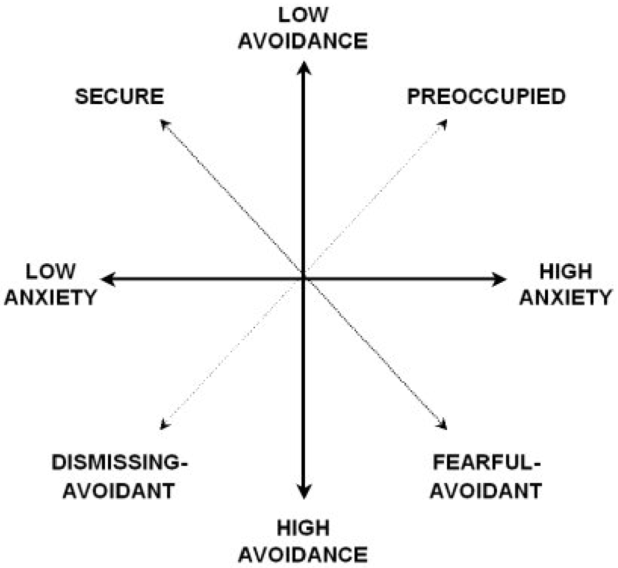


Figure 4.1. The four styles (secure, preoccupied/anxious-ambivalent, fearful-avoidant and dismissive-avoidant) and two dimensions (anxiety and avoidance) which the results of the ECR measure of adult attachment conceptually fall on.

et al. (2010)). This measures the hyperactivation and deactivation of the caregiving system (Figure 4.2).

As attachment styles are conceived of within a continuum by the ECR, so too are caregiving styles conceived of within the CSS. Hyperactivation and deactivation are orthogonal (much as attachment anxiety and attachment avoidance are), and combine to form a two-dimensional space within which individual differences in caregiving orientation are represented. Low scores on each dimension results in optimal caregiving. High scores in one dimension and low in the other, results in either an anxious-hyperactivated caregiving style or an avoidant-deactivated caregiving style. Whilst, fi-

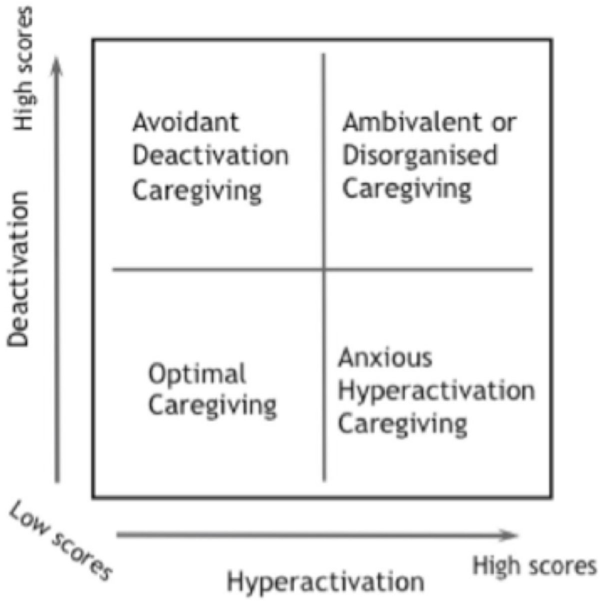


Figure 4.2. The Four-Dimension Caregiving System Scale (CSS) Model.

nally, high scores in both dimensions results in ambivalent or disorganised caregiving.[3]

Behaviourally this translates as the following: Hyperactivated caregiving styles are behaviourally intrusive, employing clinging, controlling and coercive behaviour, and using cognitive and behavioural efforts to establish proximity with others. Conversely, deactivated caregiving styles are categorised by the suppression or inhibition of proximity seeking desires and actions, resulting in the maintenance of physical and emotional distance from others, and a discomfort with interdependence and intimacy.

---

[3] Note that the existence of this fourth style – ambivalent/disorganised – exemplifies the importance of not aggregating data on a continuum, there are variations and combinations.

The CSS has not been used in any AAT or HRI studies to date. However, as was discussed in Chapter 3 the general concept of attachment and its related measures have been used to help understand a variety of different human-other interactions. For example, attachment style metrics have been used in AAT in order to explore potential variation in patient responses to animal interventions, despite the fact that the original notion of attachment was developed to describe the bond between an infant and its primary human caregiver. Barker et al. (2015) explored the effect of AAT on anxiety and pain in hospitalised children. As in the Shiloh et al. (2003) experimental design employed in the present study reported in this chapter, pre- and post-intervention questionnaires were administered. In order to explore whether a variable related to the child itself was moderating the outcome of the intervention the Attachment Questionnaire for Children (Sharpe et al., 1998) was also administered. Of particular relevance to Barker et al. (2015) was the concept of a safe haven (see Chapter 3, Section 3.3.1) to which a child can turn when exposed to threat. Animals have been shown to provide attachment security (Beck and Madresh, 2008) and may be able to fulfil attachment functions such as proximity maintenance as well (Kurdek, 2008). However, the impact of attachment style on pain and anxiety outcomes in individual's undertaking AAT had not yet been addressed prior to the Barker and colleagues study (2015). The results provided some support for attachment style's role as a potential moderating variable on the effect of the AAT on pain and anxiety outcomes: Children with a secure attachment style reported lower pain and anxiety at baseline, with large effect sizes for differences in both variables too. Although these results are not overwhelming they do point to the usefulness of HHI metrics within human-other interaction domains.

It should also be noted that AAT therapists take client personality and individual differences into account when selecting the right animal for their treatment (Parshall, 2003), although no standardised guidelines exist outlining how this should ideally be done. However, attachment theory has been used as a framework for exploring human-animal relationships due to the interpersonal nature of the dyad. For example, Zilcha-Mano et al. (2011) reported that individual differences in attachment do occur in the anxiety and avoidance domains with respect to a person's relationship with their pet. However, prior to the present study the impact of individual differences in attachment style upon the outcome of a human-PARO interaction had not been examined.

Taken together previous literature such as that described above supports the decision in the present study to include individual differences' metrics in caregiving and attachment in an exploratory manner, to look for mediating effects potentially involved in other observed differences in FS and physical engagement.

In sum, research into individual differences in caregiving and attachment styles indicate that individuals adopt certain behavioural strategies to regulate their interpersonal boundaries. Thus, if individuals with deactivated caregiving (or avoidant attachment) styles have a tendency to deactivate emotional responses, whilst individuals with hyperactivated caregiving (or anxious attachment) styles have a tendency to hyperactivate emotional responses, it can be hypothesised that the same behavioural patterns may be witnessed during an interaction with PARO. This may especially be the case given PARO's design as a RAT agent intended to elicit caregiving behaviours from its user, in a manner similar to a doll being used in doll-therapy (see Chapter 2, Section 2.4.1).

With respect to the current study translating these known behavioural differences, in caregiving styles in particular, into observable differences in physical engagement with PARO via touching behaviours could be expected to result in trends such as those in Figure 4.3.

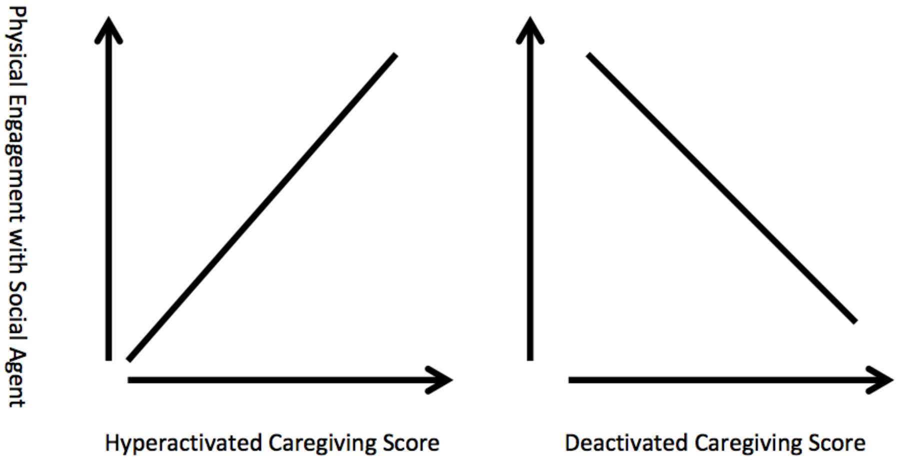


Figure 4.3. (Left) The higher a hyperactivated caregiving score an individual has the more likely they are to physically engage a social agent. (Right) The higher a deactivated score an individual has the less likely they are to physically engage a social agent.

4.2.3 Hypotheses

The primary aim of the work reported in this chapter was to explore whether an interaction period with the biomimetic PARO therapeutic robot would lead to a positive change in participant's felt security (FS) as measured with the Felt Security Scale (FSS; Luke et al. (2012)).

This study also sought to discover *why* PARO might be having a positive impact on its user's FS. Two potential mechanisms of such an effect were therefore explored.

The first hypothesis was that participants with lower pre-interaction FS scores would physically engage PARO less during the interaction period than participants with higher pre-interaction FS scores. It was further hypothesised that participants who spent the majority of their time in the interaction with PARO physically engaging the robot via observably proximate touching behaviours (such as caressive strokes) would have greater FS delta scores than participants who spent the majority of their interaction time with PARO either not physically engaging the robot at all, or physically engaging the robot via touching behaviours that were more distal (such as with exploratory pokes).

These physical engagement analyses were conducted by first coding video recordings of each participant's interaction period in accordance with a physical engagement scheme developed for this thesis, as no existing comparable scheme exists. Categorisation of touching behaviours evolved from a delineation between caressive stroking (proximate) and exploratory poking (distal) behaviours in the pilot study, to a detailed continuum coding scheme developed in the main study which graded observed behaviours based on overall physical proximity to the robot.

The second hypothesis was that physical engagement with the robot as measured via touching behaviours would be mediated by the participant's caregiving and attachment styles. Thus, it was predicted that a participant who scored more highly on deactivated or avoidant styles would be observed to be spending most of their time in the interaction period either not touching the robot or touching the robot in an distal manner (by touching the robot in, for example, a more exploratory rather than caressive manner), and subsequently those participants would have smaller FS delta scores. Conversely, participants who scored more highly on hyperactivated or anxious styles would be more inclined to spend most of their time touching the robot in a more proximate manner (for example, by employing caressive strokes), and subsequently those participants would have greater FS delta scores.

4.3 The Robotic Platform: PARO

The social biomimetic robot platform, PARO (Shibata et al., 2001) was chosen for this study. PARO is a therapeutic device modelled on a Canadian harp seal pup (Figure 4.4). The device, currently in its eighth iteration, is covered in white artificial fur beneath which are ten tactile sensors. PARO also has light, audition, temperature and gyroscopic sensors. It has eight actuators; two on upper and lower eyelids, one for rotation of eyes, two for the neck, one for each front flipper, and one for its tail. PARO weighs 2.8kg.

Figure 4.4. PARO's *cute* design is intended to elicit positive responses from humans.

PARO's primary use is in therapy sessions attended by individuals suffering from dementia and other conditions of cognitive decline. Here, PARO is utilised as a robotic replacement for domesticated animals otherwise used in AAT. PARO engages its user with basic capabilities: sensing touch, recognising a limited amount of speech, expressing small utterances and moving its head, flippers and tail. PARO does not locomote, and is designed to be held and fussed over. PARO is an excellent example of a biomimetic robot design. The relevance of biomimetics to human-robot interactions, more generally, is widely attested. Robots that are biomimetic in their morphology, in the way they move, and that have expressive faces are immediately and intuitively engaging, owing to humans' familiarity with mammalian channels for conveying emotion and intent (Breazeal and Scassellati, 1999). Indeed PARO works, in evoking the intended response from naïve subjects, because of this familiarity humans share with biological communication channels. Whilst the lack of *a priori* knowledge of seals, more generally, aids in positive outcomes with PARO: preconceived ideas of the capabilities of familiar animals, such as domestic cats, have been shown to have a negative impact on users who compare such morphologically similar robot's capabilities with the superior capabilities of their living equivalent (Shibata et al., 2001). For the purposes of this study, in exploring the effectiveness of biomimetic robot platforms, PARO is a good choice.

4.4 Pilot

The primary aim of the study reported in this chapter was to explore whether an interaction period with PARO would lead to a positive change in participant's FS. Further, the study also sought to discover *why* PARO might be having a positive impact on its user's FS. One mechanism employed to explore this was based on behavioural observations, and was intended to discover whether a participant's FS score pre-interaction was indicative of how that participant physically engaged with PARO during an interaction period, and also whether the way in which a participant physically engaged PARO impacted their FS delta score. This required the development of a new coding scheme aimed at producing a numerical measure of physical engagement based on existing understandings of the impact of proximate touching behaviours observed between humans and animals in AAT: behaviours best exemplified by stroking actions.

The second mechanism hypothesised that physical engagement with the robot as measured via touching behaviours would be mediated by the participant's caregiving and attachment styles, such that in turn these individual differences would have an impact on a participant's FS delta score.

Given the novelty of this study's design a pilot was run before the main study. This served to not only check the relevance of the FSS metric to an HRI context, but also to explore the idea that caregiving and attachment styles can be linked to observed differences in physical engagement as measured via touching behaviours in an HRI context. The pilot study also served to help develop a coding scheme for measuring touching behaviour.

Two questions were therefore addressed by this pilot before the main study took place. Firstly, are an individual's FS scores related to how they physically engage with a PARO? And secondly, do individual differences in caregiving and attachment styles relate to how a participant physically engages with a PARO?

4.4.1 Method

Ethical approval for this study was granted by the Ethics Committee in the Department of Psychology at The University Of Sheffield.

Participants

Participants (n = 10, 3 female; M age = 23.7, SD = 3.59) were recruited informally from The University Of Sheffield. All participants were healthy, with no known physical, auditory or visual impairment. Prior to study participation written informed consent was obtained from each participant. Participants were compensated for their time with a one-off payment of 5GBP. There were no instances of withdrawal.

Pre-session measures of individual differences

Caregiving style was measured using the 20-item Caregiving Structures Questionnaire (CSS; Shaver et al. (2010)). The metric produces two scores, one of hyperactive caregiving and one of deactivated caregiving, which combine to produce a general trait pattern of adult caregiving style for a participant which fall into one of four caregiving styles: optimal caregiving, avoidant-deactivated caregiving, anxious-hyperactivated caregiving and ambivalent/disorganised caregiving. Participants responded to this measure on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Representative items include *It's hard for me to work up much interest in helping others* and *I sometimes feel that I intrude too much while trying to help others*. Evidence for the reliability and construct validity of the CSS is rigorously made by Shaver et al. (2010).

Attachment style was measured using the 12-item Experiences in Close Relationships Scale - Short Form (ECR-S; Wei et al. (2007)), which is derived from the longer, 36-item ECR (Brennan et al., 1998). Results from the ECR-S produce two scores, one of attachment anxiety and one of attachment avoidance, which combine to form four attachment styles: secure, preoccupied, dismissive-avoidant and fearful-avoidant. An individual's attachment style can be used to help indicate how that individual will likely behave in relationships with another. Participants responded to this measure on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Representative items include *I need a lot of reassurance that I am loved by my partner* and *It helps to turn to my romantic partner in times of need* (reverse scored). Evidence for the reliability and construct validity of the ECR-S is rigorously made by Wei et al. (2007).

Fantasy disposition was also measured using the 7-item Fantasy Scale from the 26-item Interpersonal Reactivity Index (IRI; Davis et al. (1980)).

The Fantasy Scale measures the tendency for a person to identify with characters in movies, novels, plays and other fictional situations. This scale was included in order to control for participant individual differences in tendencies to transpose themselves imaginatively into the feelings and actions of fictitious characters, such as a non-living agent like PARO. Participants responded to this measure on a Likert scale (0 = *Does not describe me well*, 4 = *Describes me very well*). Representative items include *I really get involved with the feelings of the characters in a novel*, and *Becoming extremely involved in a good book or movie is somewhat rare for me* (reverse scored). An extensive body of research supports the reliable psychometric properties and construct validity of the IRI (2830 citations as of 21/04/2016).

Outcome measure: Delta Felt Security

The key dependent variable serving as an interaction impact check for this study is a felt security change score, called the delta score, measured by the 16-item Felt Security Scale (FSS; Luke et al. (2012)) which was taken pre- and post-interaction with PARO. The delta score is obtained by subtracting a participant's FSS pre-interaction score from their FSS post-interaction score. The FSS administered in this study is presented as a visual analogue scale (VAS) where each question is answered by making a mark along a 100mm line on either end of which is an item relating to the full dimensionality of felt security being measured by the FSS (see Figure A.1, in Appendix A). Participants indicated how *comforted, supported, looked after, cared for, secure, safe, protected, unthreatened, better about myself, valued, more positive about myself, I really like myself, loved, cherished, treasured, and adored* they were currently feeling. Total scores for the felt security items were computed by Luke et al. (2012), who concluded that the items form a reliable scale (Cronbach's $\alpha = .97$; $M = 4.25$, $SD = 1.39$).

Procedure

After volunteering to participate in the study participants were sent an email containing a link to an online questionnaire which had to be completed prior to attending the interaction session. The questionnaire was in the form of a Google Form entitled *Exploration of HRI lab parameters with a PARO robot - pre-interaction*, which opened with a statement of consent.[4] The questionnaire included questions about participant's demographics (age, gender, nationality); the ECR-S (Wei et al., 2007); CSS (Shaver et al., 2010); and the Fantasy Scale questions from the IRI (Davis et al., 1980). Once the completed questionnaire was received back by the experimenter the participant was sent a second email inviting them to

---

[4] Opening consent statement on the pre-session online questionnaire: *In completing and submitting this questionnaire you are indicating that you agree with the following statement: 'I understand that the purpose of this research is to find out more about experimental methodology in human-robot interaction. I understand that submitting this questionnaire indicates my informed consent to my data being used for this purpose. I understand that I do not have to take part and am free to withdraw (by not completing the questionnaire) at any point up until submitting the questionnaire via the final 'submit' button.'*

attend a robot-interaction session at a convenient time.

The interaction session took place in a private office in the Department of Computer Science at The University Of Sheffield. Upon arrival participants were seated in the foyer area of the department and given, 1) information sheet number one which explained the study, whilst not mentioning that the interaction would be recorded; 2) the first written consent form to sign, and; 3) a pre-session Felt Security Scale (Luke et al., 2012) form to complete. Whilst the participant attended to this paperwork the experimenter went to the private office and turned on the PARO robot and a hidden Replay XD720 Camera that was concealed amongst books on a shelf so as to go unnoticed by participants. This was to ensure that natural behaviours were captured which would not be compromised by participants potentially feeling nervous about being watched.

The experimenter then returned to the participant. After taking the participant's paperwork and asking them if they had any questions the experimenter led them to the private office. Outside the room, whilst the door was still closed, the experimenter gave each participant the same scripted instruction.[5]

---

[5] *Please can you turn your phone off. Please enter the room and sit on the blue chair. PARO is in front of you. Please interact with the robot as you wish. Pick it up, carry it, explore its features, you may walk around with the robot if you wish. I will come back after five minutes and knock the door and ask you to fill in the form in front of you. When you are done please let me in. Then we will talk about your experience. I will not be coming in with you as I do not want to influence how you interact with PARO in anyway, afterwards during your interview, I am looking for your complete unbiased opinion of PARO. Is that all right? Do you have any questions?*

Inside the room a table had been arranged in the centre with a blue chair to one side of it. An unused desk and a bookcase were at the far end of the room. The camera was hidden in the bookcase. PARO was positioned facing the chair on the table. Face down on the table was a second FSS form.

After five minutes the experimenter returned to the room, knocked on the door, and waited for the participant to grant them entry to the room. The experimenter then opened the door and asked the participant to turn over the piece of paper on the table and complete an FSS form for a second time. The experimenter then left the room and asked that the participant open the door to let them back in once the FSS form was complete. At that time the experimenter took a seat at the table with PARO on. PARO was switched off and moved to one side. The experimenter then explained to the participant that they had been covertly video recorded, and turned off the hidden camera. All participants were then given an opportunity to have their recording deleted if they wished to withdraw from the study at that time. Participants were informed that they would still receive monetary compensation if they so wished to withdraw.

The experimenter then gave the participant a second information sheet, which explained that the study was recorded in order to take behavioural observations, and a second written consent form to sign which explained that video data would be analysed as well as the questionnaire data. The experimenter then debriefed the participant fully with these two forms, handed the participant their monetary compensation, and finally led the participant out of the department.

Behavioural observations

Video data was coded on a second-by-second basis for measures of physical engagement with PARO using The Observer$^{®}$ XT software. Five touching behaviours were coded for: *no touch; whole palm touch; fingertip touch; other touch; and hit.*

Of the five types of touching behaviours coded for three were focussed on: *no touch; whole palm touch; and fingertip touch.* Whole palm touches indicated the most proximate physical engagements observed with PARO (see Figure 4.5) and were deemed representative of high emotional engagement via touches that established close proximity with the robot. This decision was based on existing literature in AAT on the positive impact of stroking animals (see Section 4.2.1). Fingertip touches indicated a more distal physical engagement with PARO, and were exemplified by poking behaviours (see Figure 4.5). These fingertip touches were deemed representative of low emotional engagement with PARO via touches that maintained physical and emotional distance with the robot. This decision was based on an understanding of deactivated and avoidant behavioural styles which are often conceptualised as those intent on maintaining physical and emotional distance from other social agents (see Section 4.2.2).

Hits, no touching observed, and other touches (i.e. any touch that did not belong to the other four categories, e.g., picking PARO up to turn it over) were also coded for. Thus a composite physical engagement ratio was calculated for each participant for each of the touches they engaged in with PARO during the interaction period, see Figure 4.6 and Table 4.1.

Figure 4.5. Left panel: Proximate physical engagement, such as stroking motions, were coded as *whole palm touches*, wherein a participant was observed as emotionally and physically close to the robot. Right panel: Distal physical engagements, such as poking PARO, were coded as *fingertip touches* wherein a participant was observed as emotionally and physically distant to the robot.

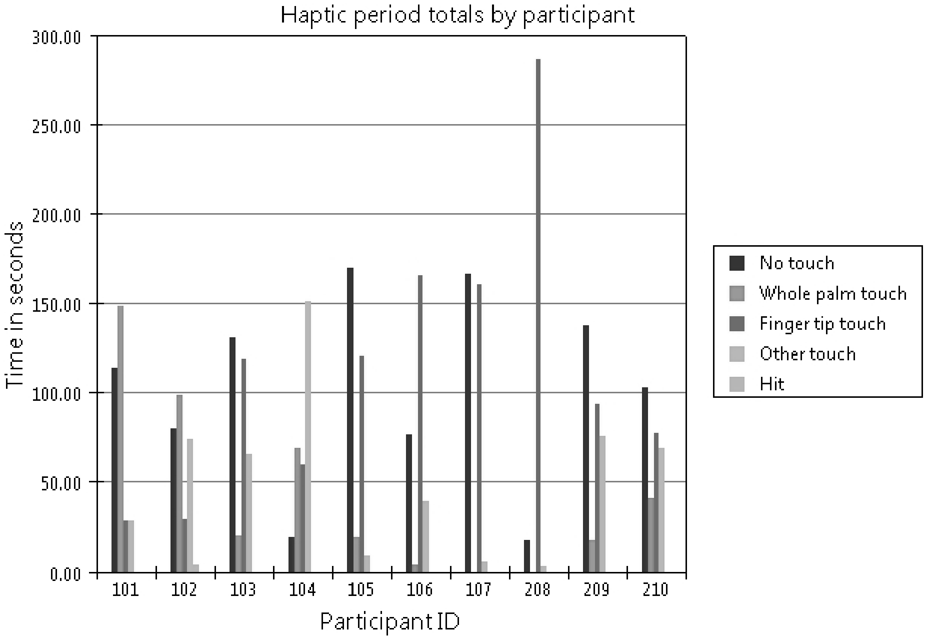


Figure 4.6. Graph of total duration in seconds calculated for each participant by five haptic behaviours.

Table 4.1. Total duration in seconds calculated for each participant by five haptic behaviours.

Participant ID	No touch	Whole palm	Fingertip	Other	Hit	Duration
101	114.16	149.23	28.94	28.81	0	321.14
102	80.18	98.96	29.51	74.24	4.23	287.13
103	130.99	20.36	119.18	66.00	0	336.53
104	19.75	69.82	59.93	151.82	0	301.32
105	170.19	19.18	121.05	9.39	0	319.82
106	77.20	3.92	166.13	40.24	0	287.50
107	167.12	0	161.00	6.37	0	334.48
208	17.57	0	286.57	3.12	0	307.26
209	138.13	18.18	94.13	76.59	0	327.03
210	102.92	41.82	78.14	69.18	0	292.06

Reliabilities

All measures of physical engagement were checked by the author. Due to the small sample size the physical engagement coding scheme employed in the pilot was not second coded. This decision was also made due to the large changes that occurred between the coding scheme used in the pilot and that finally developed for use in the main study (see Section 4.5.1, for a full discussion of the final coding scheme).

Reliability analyses were run on each subscale of each of the pre-session measures of individual differences. All subscales had high reliabilities: CSS deactivated subscale, Cronbach's α = .890; CSS hyperactivated subscale, Cronbach's α = .910; ECR-S avoidant subscale, Cronbach's α = .864; ECR-S anxious subscale, Cronbach's α = .922; and the Fantasy Scale, subscale of the IRI, Cronbach's α = .800. Total scores for the pre-interaction FSS items were also computed. The scale had high reliability, Cronbach's α = .883.

Statistical analyses

For each participant their total time spent engaging with PARO in each of the five behaviours coded for, *no touch; whole palm touch; fingertip touch; other touch; and hit*, were normalised against each participant's total time spent in the interaction phase. Hit was subsequently excluded from analysis as only one participant hit PARO for a total duration of 4.23 seconds.

All statistical analyses were run in IBM SPSS Statistics 23 for MAC OS X (10.10.5). All plots were created using a custom script produced in MATLAB R2014b for MAC OS X (10.10.5).

Note that the Fantasy Scale was not included in analysis. It had initially been intended to be used as a controlling variable in the event that the caregiving and attachment style data was correlated with touch-type duration. As all those relationships were non-significant the Fantasy Scale was no longer required.

4.4.2 Results

In order to establish whether an interaction period spent with PARO positively impacted participants' FS a dependent t-test (with a 95% confidence interval) was conducted comparing pre- and post-interaction FS scores. The t-test revealed that on average, participants' (n = 10) pre-interaction FS score (M = 77.28, SE = 3.22) was not significantly different to participants' post-interaction FS score (M = 79.56, SE = 3.56), $t(9)$ = -0.605, p = .560. However, eyeballing of the data revealed one clear outlier, participant three: FS delta score = -27.06. This value was over two standard deviations away from the mean (FS delta score SD = 11.91) and as such was subsequently excluded from all analyses (see Table 4.2 for all values).[6] The dependent t-test was re-run. Results from this second t-test, with a 95% confidence interval, reveal that on average, participants' (n = 9) pre-interaction FS score (M = 76.25, SE = 3.42) was significantly different to participants' post-interaction FS score (M = 81.78, SE = 3.11), $t(8)$ = -2.620, p = .031. The positive trend, with post-interaction FS scores being on average higher than pre-interaction FS scores, reveal a positive impact of the interaction period with PARO on participants' FS scores.

Table 4.2. FS pre- and post-interaction and delta scores for each participant.

Participant ID	FS Pre-Interaction	FS Post-Interaction	FS Delta Score
101	69.19	79.84	10.66
102	72.53	84.97	12.44
103	*86.56*	*59.50*	*-27.06*
104	71.97	85.59	13.63
105	91.50	93.75	2.25
106	88.44	94.63	6.19
107	79.69	82.22	2.53
208	82.91	77.34	-5.56
209	70.81	70.59	-0.22
210	59.19	67.13	7.94

[1] Participant 103 italicised, note the large FS delta score. Participant 103 subsequently excluded from dependent t-test analysis.

Note that the sample was too small to look for effects of gender or age.

---

[6] Due to the small sample size tests for normality were not conducted.

Scatterplots, with lines of best fit, visualising the relationship between FS pre-interaction (where higher scores = higher feelings of felt security in participants ($n = 9$) immediately prior to interaction with PARO) and the three types of touching behaviours of most interest (*no touch; whole palm touch; and fingertip touch*) were produced in order to explore whether a participant's pre-interaction FS score was indicative of how participants physically engaged with PARO during their robot-interaction session (Figure 4.7).

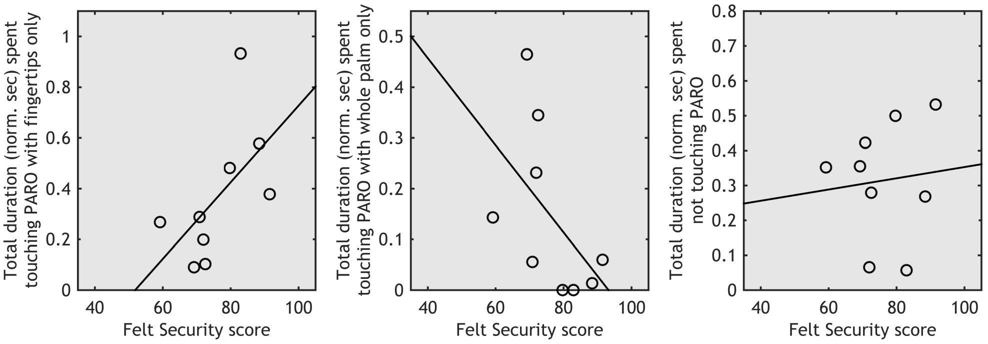


Figure 4.7. Scatterplots ($n = 9$) with lines of best fit visualising the relationship between FS pre-interaction and three types of touching behaviours: fingertip touches (left), whole palm touches (centre), and no touch (right).

Trends indicate that participants pre-interaction FS score correlated positively with the proportion of normalised total time spent touching PARO with fingertips or not touching PARO at all. Conversely, pre-interaction FS scores were negatively correlated with time engaged in whole palm touching behaviour, with higher pre-interaction FS scores associated with less time spent physically engaging PARO with behaviours such as stroking.

To explore whether there was a relationship between how a participant

($n = 9$) physically engaged with PARO during the interaction period and their FS delta score, simple linear regressions were calculated to predict FS delta score (dependent variable) based on each of the four normalised touch-type duration values (independent variables: *no touch; whole palm touch; fingertip touch; and other touch* (*hit* excluded, see section *Statistical analyses* above)). (Table 4.3.)

Table 4.3. Simple linear regressions ($n = 9$) predicting FS delta score (DV) based on each of the four normalised touch-type duration values (IVs).

Normalised Touch-Type	FS Delta Score F	FS Delta Score p
No touch	0.10	.76
Whole palm touch	9.67	.02*
Fingertip touch	13.04	.01*
Other touch	5.49	.05

[1] F ratio and p value for four remaining normalised touch-type duration values, excluding *hit*.

[2] *Denotes significant effect.

A significant regression equation was found on FS delta score of whole palm touches, $F(1, 7) = 9.67$, $p = .017$, with an R^2 of .58, and on FS delta score of fingertip touches, $F(1, 7) = 13.04$, $p = .009$, with an R^2 of .65. These relationships are visualised in scatterplots (with lines of best fit) in Figure 4.8.

In order to explore whether there was a relationship between participants' individual differences in caregiving and attachment styles and how a par-

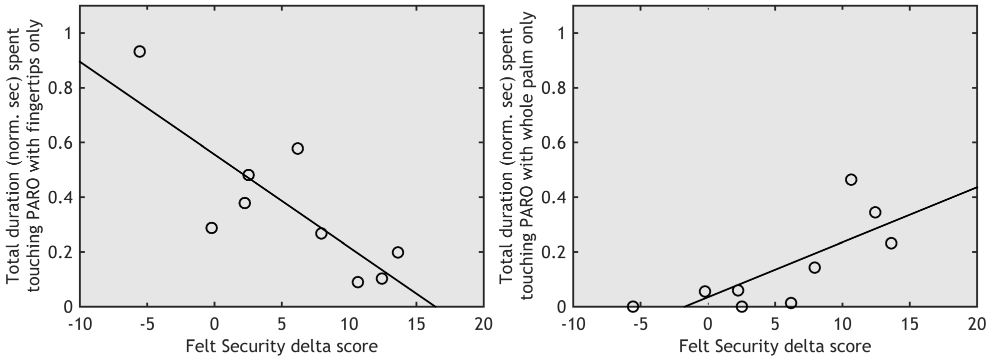


Figure 4.8. Scatterplots with lines of best fit visualising the relationship between FS delta score ($n = 9$) and physical engagement with PARO coded as fingertip touches (left) and whole palm touches (right).

ticipant physically engaged with PARO during the interaction period, simple linear regressions were calculated ($n = 9$) to predict amount of time (normalised) spent engaged in the four touch-type behaviours with PARO (*no touch; whole palm touch; fingertip touch; and other touch*: dependent variables) based on each subscale from the pre-session measures of individual differences (*ECR-S attachment avoidance; ECR-S attachment anxiety; CSS deactivated caregiving; and CSS hyperactivated caregiving*: independent variables). All regression equations were non-significant.

Non-significant results were unsurprising due to the small sample size of the pilot. However, further data visualisations focussed on whole palm touch and fingertip touch - those physical engagements shown to have significant relationships with FS delta scores in the previous simple linear regressions - were conducted to look for trends and patterns that could help inform the main study.

Scatterplots of CSS and ECR-S scores were created which visualised participants' ($n = 9$) caregiving and attachment styles, by mapping the scores over the caregiving and attachment style quadrants which are created when deactivated or avoidant scores are combined with hyperactivated or anxious scores to create a total caregiving or attachment style score for each participant (see Figure 4.9 for CSS scores, and Figure 4.10 for ECR-S scores). Scatterplots graph participants who were engaging in high or low amounts of fingertip or whole palm touching behaviours as normalised against total interaction time. High and low values were arbitrarily designated by dividing the sample equally: the five highest values separated from the five lowest values (after which participant 103 was excluded from the graph).

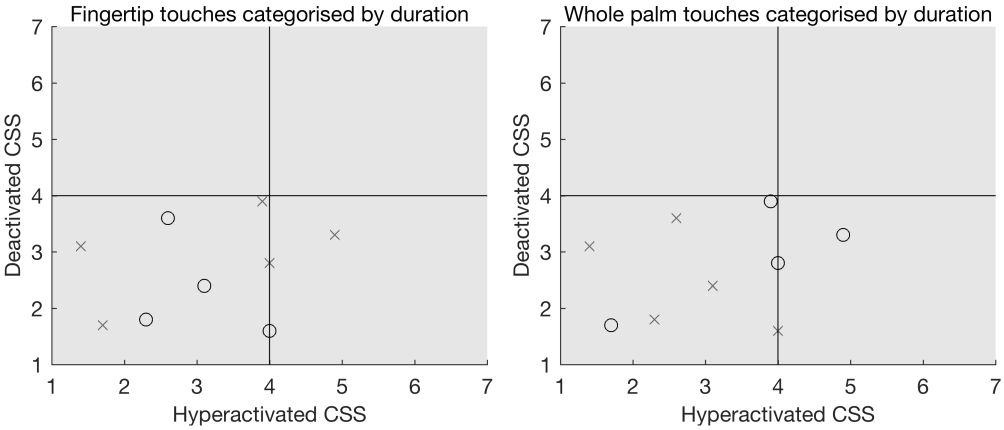


Figure 4.9. Scatterplots ($n = 9$) of relationship between CSS scores and fingertip (left) and whole palm (right) touch duration (high (X) or low (O): normalised against total interaction time in seconds). High denoted with X; Low denoted with O.

Note that these graphical representations of participants' locations on the caregiving and attachment style dimensions have x and y axis values that are equivalent to the range of possible scores on the CSS and ECR-S questionnaires (Likert values of 1-7). Further, note that the majority of participants fall in the secure/optimal style, bottom left quadrant (indicating low scores in each dimension). The normality of the population meant that there were no extremes to the data, however there was still a distribution in this sub area of each of the two-dimensional spaces. It was assumed that the larger population of the main study might result in a wider distribution of data across the four quadrants of each of the CSS and ECR-S scores.

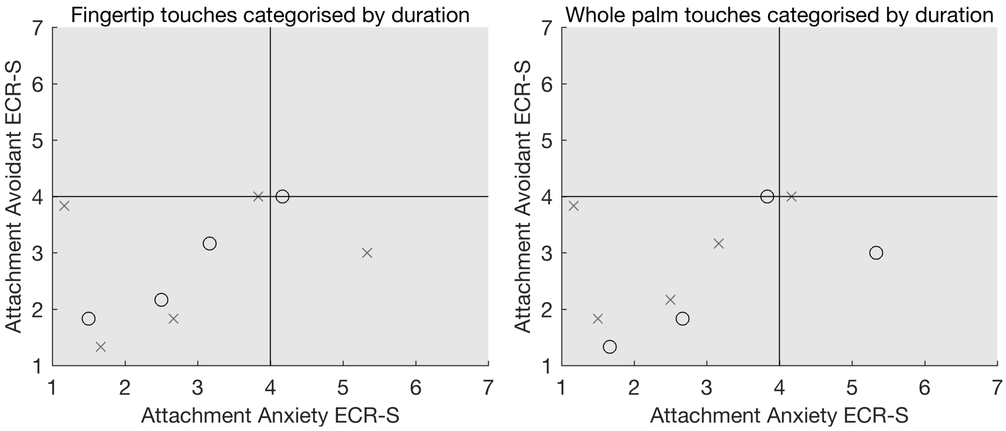


Figure 4.10. Scatterplots ($n = 9$) of relationship between ECR-S scores and fingertip (left) and whole palm (right) touch duration (high (X) or low (O): normalised against total interaction time in seconds). High denoted with X; Low denoted with O.

4.4.3 Discussion

After removing an outlier a significant difference between pre- and post-interaction FS scores was found, with an overall positive trend in the data such that post-interaction scores were more likely to be higher than pre-interaction scores indicating a positive effect of the interaction period with PARO. This finding supports the primary aim of the chapter which was to see if interacting with PARO would lead to a measurable positive impact on its user's FS. In attempting to answer why this might be happening this pilot addressed two questions. Firstly, are an individual's FS scores related to how they physically engage with a PARO? And secondly, do individual differences in caregiving and attachment styles relate to how a participant physically engages with a PARO?

In answering the first question, results indicate that individuals with higher FS pre-interaction scores were more likely to spend the majority of their time interacting with PARO by physically engaging the robot with their fingertips or not touching PARO at all. Whilst individuals with lower FS pre-interaction scores were more likely to spend the majority of their time with PARO physically engaging it with their whole palm. Further, individuals who spent the majority of their interaction time touching PARO with their fingertips had smaller FS delta scores than individuals who interacted with PARO with their whole palms, who had larger FS delta scores. Thus it appears in this small sample that an individual's baseline FS score does impact how they interact with PARO, and subsequently how an individual interacts with PARO does have an impact on their FS delta score.

This result must be considered not only in light of the small sample but also whilst considering the FS score ceiling effect. That is, the higher an individual's pre-interaction FS score, the smaller their potential FS delta can be. Conversely, the lower an individual's pre-interaction FS score, the larger their potential FS delta score can be.

In answering the second question, results from this pilot show that individual differences in caregiving and attachment styles do not relate to how a participant physically engages with a PARO. Although the results were graphically represented in order to look for trends and patterns to help inform the main study, no clear patterns were found. However, in Figure 4.9, right, one pattern of note can be seen. Here whole palm, that is proximate physical engagement behaviour, is mapped onto participants' CSS scores. Individual's in the optimal caregiving quadrant display high duration whole palm touches, whilst those in the anxious-hyperactivating caregiving quadrant display low duration whole palm touches. This visualisation contradicts the hypothesis that individuals scoring highly on anxious-hyperactive caregiving would be more likely to engage in clinging, proximate behaviours with other social agents (Figure 4.9).[7]

The pilot also served to help develop a coding scheme for the main study, as no existing scheme exists with which to examine physical engagement with a PARO robot. Although the scheme developed in the pilot served its purpose in this small sample, some issues were found when it came to applying it to the larger main study.

---

[7] The scatterplots of relationship between ECR-S scores, and fingertip and whole palm touch durations (Figure 4.10) are more ambiguous, so no extrapolations were made.

The simple division of stroking (proximate) versus poking (distal) behaviours was not comprehensive enough to capture the full dimensionality of what was being witnessed during robot interactions. This discovery occurred after coding half of the main study videos with the coding scheme developed for the pilot. How a participant touched PARO was not directly related to the intention behind the touch such that it could be objectively observed. For example, sometimes a fingertip touch looked more like a *caressive* fingertip stroke than an *exploratory* fingertip poke. However, according to the coding scheme developed in the pilot, if such an ambiguous fingertip touch were coded as *fingertip touch* it would be ultimately categorised as a distal touch and not a proximate touch. Given this revelation a new scheme was developed midway through the analysis of the main study results. This new scheme would produce a series of interaction scores based on a graded system of behaviours categorised by a participant's physical proximity to PARO. This better captured the activity of the participant as the interaction scores produced a measure of how physically engaged a person was with the robot (with higher scores indicating closer and more intimate physical interaction). This new scheme will be described in detail in Section 4.5.1.

Overall the pilot study indicated that this experimental design worked well, and the results, though mostly non-significant, were interesting enough to be pursued in a larger study.

4.5 Control Study

In order to strengthen the findings of the subsequent main study, a control experiment was conducted to test whether a change in felt security scores would be found without a robot interaction period. Participants (n = 30, 16 female) were recruited formally via The University Of Sheffield student volunteers email list. All participants were healthy adults from whom informed consent was obtained prior to the experimental phase. Participants were compensated for their time with a one-off payment of 5GBP.

The control experiment took place in Sheffield Robotics at The University Of Sheffield. Upon arrival participants were seated outside the HRI lab and asked to complete ethics forms as well as a pre-session FSS (Luke et al., 2012). They were then seated inside the HRI lab and left alone for 5 minutes. At that time the experimenter returned and asked the participant to once again complete an FSS. Participants were told this was not because their previous answers had been incorrect, but that it was to see how they felt now they were in the lab itself. Once that was completed the participants were left alone for five minutes to interact with a PARO and a MIRO (see Chapter 5). After that interaction participants were interviewed by the experimenter in order to obtain information about the two platforms to help with the design of future studies.

A paired sample t-test revealed that on average there was no significant effect of waiting alone in the HRI lab on pre-interaction FS scores ($M = 77.66$, $SE = 2.47$) compared to post-interaction scores ($M = 77.02$, $SE = 2.58$), $t(29) = 0.673$, $p = .506$.

The descriptive data from this control is included in Table 4.6 for comparison.

4.6 Main Study

Given the pilot study and the results described above the main study was pursued following an almost identical experimental design. The largest change between the pilot and main studies were those made to the coding scheme. In the main study a new coding scheme was developed to produce an intimacy score which better conceptualised how proximate or distal a participant was with PARO during the interaction period.

The hypotheses for the main study remained those outlined in Section 4.3.

4.6.1 Method

Ethical approval for this study was granted by the Ethics Committee in the Department of Psychology at The University Of Sheffield.

Participants

Participants (n = 61, 32 female; M age = 24.42, SD = 3.99) were recruited formally via The University Of Sheffield student volunteers email list. All participants were healthy, with no known physical, auditory or visual impairment. Prior to study participation written informed consent was obtained from each participant. Participants were compensated for their time with a one-off payment of 5GBP. Due to mechanical failure participant number 58 was removed from analysis and replaced with a new participant, number 61 (both female).

Measures

The pre-session measures of individual differences in attachment style, caregiving style and fantasy, as well as the outcome measure of delta felt security were measured using the same metrics as in the pilot (see Section 4.4.1).

Procedure

After volunteering to participate in the study, participants were sent an email inviting them to attend an interaction session at a convenient time.

The interaction session took place in Sheffield Robotics at The University Of Sheffield. Upon arrival participants were seated alone at a computer station and directed to complete a questionnaire in the form of a Google Form entitled *Exploration of HRI lab parameters with a PARO robot - pre-interaction*.[8] The questionnaire was identical to that used in the pilot study. It contained questions on demographics, the 7-item Fantasy Scale (Davis et al., 1980), the 12-item ECR-S (Wei et al., 2007), the 20-item CSS (Shaver et al., 2010), and it opened with the same statement of consent (see section 4.4.1). Once the completed questionnaire was received into the experimenter's email box the experimenter returned to the computer station to collect the participant for the interaction session.

The experimenter led the participant to the Human-Robot Interaction (HRI) lab at Sheffield Robotics. The participant was seated outside the room and given, 1) information sheet number one which explained the study whilst not mentioning that the interaction would be recorded; 2) the first written consent form to sign, and; 3) a pre-session FSS (Luke et al., 2012) form to complete. Whilst the participant attended to this paperwork the experimenter entered the HRI lab and turned on the PARO robot and the recording equipment. The HRI lab is fully equipped with a fixed audio-visual suite. Video recordings were captured on three Axis IP Colour Dome PTZ Cameras, with audio captured by a Sennheiser microphone. Media was recorded in real-time by Noldus Media Recorder

---

[8] This change from the pilot, to having participants complete the pre-session questionnaire on the same day as the interaction session occurred due to the speed with which participants were being recruited.

software. The cameras in the HRI lab are discrete, and it is not immediately obvious they are there. Attention was not drawn to them by the experimenter, and the fact that the interaction session was being recorded was not mentioned until after the session was over. This was to ensure that natural behaviours were captured which would not be compromised by participants potentially feeling nervous about being watched.

The experimenter then returned to the participant. After taking the participant's paperwork and asking them if they had any questions the experimenter gave each participant the same scripted instruction as had been given in the pilot (see Section 4.4.1). The participant was then let into the room and left alone to interact with PARO for five minutes. Inside the room a table had been arranged in the centre with a chair to one side of it. PARO was positioned facing the chair on the table. Face down on another table in the room was a second FSS form.

After five minutes the experimenter returned to the room, knocked on the door, and waited for the participant to grant them entry to the room. The experimenter then opened the door and asked the participant to turn over the piece of paper on the table and complete an FSS form for a second time. The experimenter then left the room and asked that the participant open the door to let them back in once the FSS form was complete.

Once the post-interaction FSS form was completed the participant let the experimenter back into the room. The experimenter took a seat at the table with PARO on. PARO was switched off and moved to one side. The experimenter then explained to the participant that they had been covertly video recorded, and turned off the camera system. All participants were then given an opportunity to have their recording deleted if they wished to withdraw from the study at that time. Participants were informed that they would still receive monetary compensation if they so wished to withdraw.

The experimenter then gave the participant a second information sheet, which explained that the study was recorded in order to take behavioural observations, and a second written consent form to sign which explained that video data would be analysed as well as the questionnaire data. The experimenter then debriefed the participant fully with these two forms, handed the participant their monetary compensation, and finally led the participant out of Sheffield Robotics.

Behavioural observations

Video data was coded on a second-by-second basis for measures of physical engagement with PARO using The Observer® XT software. Two categories were created to be coded for alongside one another: Tactile, defined as the hand observed as most engaged in an interaction with PARO, or an observed nose nuzzle which was coded for in this category regardless of hand position because a nose nuzzle represents the most intimate behaviour observed during the interaction sessions, and Position of PARO. These categories were then combined to make a 23-item classification system which encompassed all the behaviours observed in the video.

A detailed description of these categories and their coded behaviours is listed below with examples of each behaviour.

1. Tactile

   - *No touch*: Participant is not touching PARO with either their hands or their face in any way.

   - *Other touch*: Participant is manipulating PARO with their hands in a way that cannot be categorised by another other behaviour. Typically behaviours seen in this category were holding PARO in front of self, pushing PARO around the table, turning PARO onto its back.

   - *Hit*: Participant strikes PARO forcefully.

   - *Fingertip poke/hold*: Participant gently pokes PARO, or tentatively holds PARO's fore flipper between finger and thumb.

   - *Vibrissae touch*: Participant holds or strokes PARO's vibrissae between finger and thumb.

- *Whole hand touch*: Participant rests whole hand on PARO, but does not stroke PARO.
- *Fingertip stroke*: Participant uses one fingertip to stroke PARO, for example under the robot's chin.
- *Whole hand stroke*: Participant uses whole hand to stroke PARO. Typically a long movement running along PARO's body.
- *Nuzzle with face*: Participant uses their face to 'nuzzle' PARO. Typically participant will push their nose against the robot's.

2. Position of PARO
    - *Table*: PARO is positioned on the table in front of the participant.
    - *Held away*: PARO is being held away from the participant's body by the participant.
    - *Lap*: PARO is on participant's lap.
    - *Held against*: PARO is being held against the participant's body by the participant.

Creating a classification system based on behavioural observations

The Tactile and Position of PARO categories and their associated behaviours were then combined to create a classification system that encompassed every action seen in the videos taken during the main study. The classification system only included actions that were observed in the video. For example, *no touch* only occurred when PARO was positioned on the table, thus *no touch* is only combined with *table* in the classification system, and not with *held against*, *held away* and *lap.* Combining behaviours in this manner resulted in 23 codable actions (Table 4.4).

Table 4.4. Possible physical engagement combinations of Tactile and Position of PARO categories.

Code	Tactile	Position of PARO
0	No touch	Table
1	Other touch	Table
2	Fingertip poke/hold	Table
3	Whole hand touch	Table
4	Fingertip stroke	Table
5	Whole hand stroke	Table
6	Nuzzle with face	Table
7	Other touch	Held away
8	Fingertip poke/hold	Held away
9	Whole hand touch	Held away
10	Fingertip stroke	Held away
11	Whole hand stroke	Held away
12	Other touch	Lap
13	Fingertip poke/hold	Lap
14	Whole hand touch	Lap
15	Fingertip stroke	Lap
16	Whole hand stroke	Lap
17	Other touch	Held against
18	Fingertip poke/hold	Held against
19	Whole hand touch	Held against
20	Fingertip stroke	Held against
21	Whole hand stroke	Held against
22	Nuzzle with face	Held against

[1] When combined the Tactile and Position of PARO categories make 23 possible actions to be observed during the video recorded interaction sessions.

Note that the Tactile category had originally included two other behaviours: *hit* and *vibrissae touch*. *Hit* was only observed three times. On two of these occasions the 'hit' was better described as a 'tap' done with the fingertip. Therefore these two hits were recoded as *fingertip pokes*. The third 'hit' was actually a mini-fist bump on the PARO's nose, which was recoded as *other touch*, as it was not an aggressive hit. Thus the behaviour *hit* was removed from the classification system. All *vibrissae touches* were recoded as *fingertip pokes/holds*. This decision was made because the observed vibrissae touches were either done using the fingertip to poke the ends of each vibrissae in order to stimulate PARO to turn its head (PARO has touch capacitors in its vibrissae), or were finger and thumb grasps done to 'twizzle' the vibrissae. Given that the behaviour *fingertip poke/hold* is already coding for gentle pokes or tentative finger and thumb holds, this recoding of *vibrissae touches* to *fingertip pokes/holds* is reasonable. Further, it should be noted that in the pilot study vibrissae touches were also coded as fingertip pokes, and this attempt in the main study to break that behaviour down to extract more information was ultimately uninformative. Thus the behaviour *vibrissae touch* was removed from the classification system.

As in the pilot study the variables extracted from these coded behaviours were required to be reflective of how distal or proximate to the robot a participant was during their interaction session. Three categories were chosen which reflected this behaviour: *no touch*, *active touch* and *intimate touch*. To decide which of the 23 observable behaviour combinations would belong to each category, the behaviour combinations were ordered based on an intimacy value. Each Tactile and Position of PARO behaviour was given a value: 0-6 for each one of the seven Tactile behaviours, and 0-3 for each one of the four Position of PARO behaviours. The values were assigned based on how distal or proximate the participant and robot were to each other, given the behaviour. Thus Tactile behaviour ran from *no touch* valued at (0), to *nuzzle with face* valued at (6). Whilst Position of PARO ran *table* = (0); *held away* = (1); *lap* = (2); and *held against* = (3). When combined these two values created an intimacy value, which ran from 0-9. A second score of either (0) or (1) was then added to this intimacy value. This score was decided by the experimenter based on what had been observed in the videos. All behaviour combinations that involved minimal interaction with the robot, either because there was no touching occurring or because the observed behaviour combination was clearly explorative or tentative, received (0), whilst all other behaviour combinations received a (1) score (this second score is titled 'min/max score'). This resulted in a revised total intimacy value. Taking guidance from the label used to decide the min/max score the distribution of behaviour combinations into the three categories of *no touch*, *active touch* and *intimate touch* was then done.[9] The final series of possible observed behaviour combinations then create a continuum from no touch/PARO on the table (most distal), to nuzzle/holding PARO against the body (most proximate and intimate) (Table 4.5).

---

[9] Physical engagement codes in each of the three categories are: *No touch*, 0; *Active touch*, 1, 2, 3, 7, 8, 9, 12; *Intimate touch*, 4, 5, 6, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22.

This meant that for each participant every observable behaviour combination which made up their total interaction time with the robot could be broken down into one of three categories. The total period of time (in seconds) that each participant spent engaging in every behaviour combination corresponding to each category was then calculated. Each of these three time periods were then normalised against the participant's total interaction time in the session to create a percentage of time spent interacting with the robot in each of the three categories of *no touch*, *active touch* and *intimate touch*. In doing this the main study produced a more objective and detailed method for exploring differences in behaviours ranging from proximate to distal, than that created for the pilot.

Reliabilities

All measures of tactile and physical engagement were checked by the author. Tactile and physical engagement was second coded on 10% of the sample (6/60) by a trained research assistant naïve to the purpose of the study. The intercoder reliability tolerance window was set at four seconds. This was to accommodate variation in reaction time between coders viewing dynamic movements such as strokes. Cohen's Kappa on these scores (between the author and the trained assistant) was .54 indicating moderate agreement (Landis and Koch, 1977).

Table 4.5. Table detailing the creation of an Intimacy Value to better conceptualise the coded physical engagement behaviours on a continuum of proximate to distal.

Code	Tactile	HV	PP	PV	IV	Description	MM	RIV
0	No touch	0	Table	0	0	Nothing	0	0
1	Other touch	1	Table	0	1	Explore	0	1
2	Fingertip poke/hold	2	Table	0	2	Tentative	0	2
3	Whole hand touch	3	Table	0	3	Tentative	0	3
7	Other touch	1	Held away	1	2	Explore	0	2
8	Fingertip poke/hold	2	Held away	1	3	Explore	0	3
9	Whole hand touch	3	Held away	1	4	Explore	0	4
12	Other touch	1	Lap	2	3	Shifting	1	4
4	Fingertip stroke	4	Table	0	4	Caress	1	5
5	Whole hand stroke	5	Table	0	5	Caress	1	6
6	Nuzzle with face	6	Table	0	6	Caress	1	7
10	Fingertip stroke	4	Held away	1	5	Proximate	1	6
11	Whole hand stroke	5	Held away	1	6	Caress	1	7
13	Fingertip poke/hold	2	Lap	2	4	Proximate	1	5
14	Whole hand touch	3	Lap	2	5	Proximate	1	6
15	Fingertip stroke	4	Lap	2	6	Caress	1	7
16	Whole hand stroke	5	Lap	2	7	Caress	1	8
17	Other touch	1	Held against	3	4	Caress	1	5
18	Fingertip poke/hold	2	Held against	3	5	Caress	1	6
19	Whole hand touch	3	Held against	3	6	Caress	1	7
20	Fingertip stroke	4	Held against	3	7	Caress	1	8
21	Whole hand stroke	5	Held against	3	8	Caress	1	9
22	Nuzzle with face	6	Held against	3	9	Caress	1	10

[1] HV = Hand Value; PP = PARO position; PV = PARO Value; IV = Intimacy Value; MM = Min/Max Score; RIV = Revised Intimacy Value.

Reliability analyses were run on each subscale of each of the pre-session measures of individual differences. All subscales had high reliabilities: CSS deactivated subscale, Cronbach's $\alpha = .915$; CSS hyperactivated subscale, Cronbach's $\alpha = .883$; ECR-S avoidant subscale, Cronbach's $\alpha = .819$; ECR-S anxious subscale, Cronbach's $\alpha = .779$; and the Fantasy Scale, subscale of the IRI, Cronbach's $\alpha = .821$. Total scores for the pre-interaction FSS items were also computed. The scale had high reliability, Cronbach's $\alpha = .947$.

Statistical analyses

All statistical analyses were run in IBM SPSS Statistics 23 for MAC OS X (10.10.5). All plots were created using a custom script produced in MATLAB R2014b for MAC OS X (10.10.5). Note that the Fantasy Scale was included in a hierarchical regression analysis as a co-variate.

After calculating a set of Intimacy Values for each participant, their combined interaction time was checked against the actual total time each participant spent interacting with PARO to ensure correct calculations. All measures of individual differences were mean centred in order to standardise them.

4.6.2 Results

Table 4.6. Descriptive data (means and standard errors) of the main results for the pre- and post-interaction Felt Security Scores (FSS; DVs), for both the main (n = 60) and control studies (n = 30).

FSS Time Taken	Mean	*SE*
Pre-Interaction Main	71.69	1.82
Post-Interaction Main	77.33	1.93
Pre-Interaction Control	77.66	2.47
Post-Interaction Control	77.02	2.58

Tests for normality indicated that the pre-interaction FS scores were normally distributed, with z-skewness of -1.89 and z-kurtosis of -0.25. Results from the pre-interaction FS score Kolmogorov-Smirnov test, $D(60) = 0.090$, $p = .200$, indicate normal distribution in the sample. See Figure 4.11 for a histogram of the pre-interaction FS scores showing normal distribution. No data were excluded from subsequent analysis.

In order to establish whether an interaction period spent with PARO positively impacted participants' FS score a dependent t-test (with a 95% confidence interval) was conducted comparing pre- and post-interaction FS scores. Resulted showed that on average participants' (n = 60) pre-interaction FS score (M = 71.69, SE = 1.82) was significantly lower than participants' post-interaction FS score (M = 77.33, SE = 1.93), $t(59)$ = -4.720, p <.001 (See Table 4.6). Paired sample t-tests (each pair n = 29)

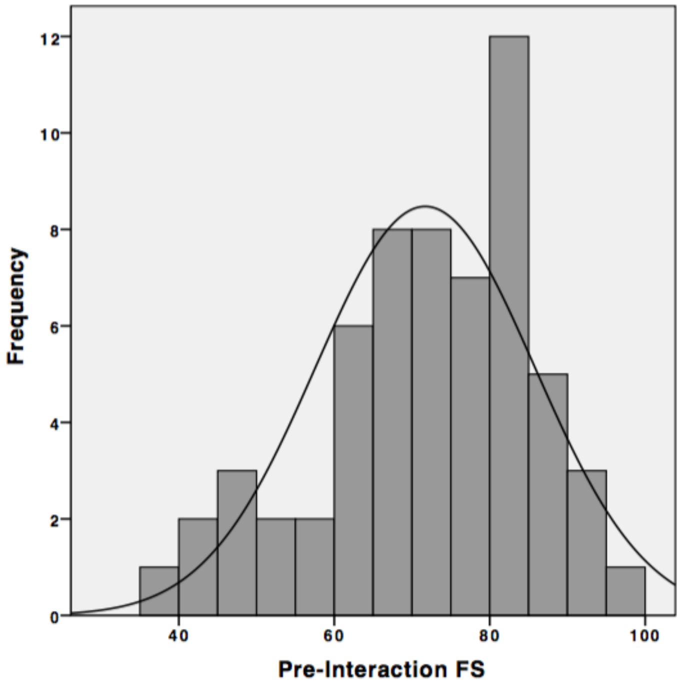


Figure 4.11. Histogram of the pre-interaction FS scores showing normal distribution.

also revealed that on average there was no significant effect of gender between male (M = 70.27, SE = 2.73) and female (M = 72.95, SE = 2.60) pre-interaction FS scores $t(28)$ = -0.678, p = .503, or male (M = 74.30, SE = 3.28) and female (M = 80.01, SE = 2.17) post-interaction scores, $t(28)$ = -1.406, p = .171.

Hypothesis 1: Relationships between physical engagement and FS scores

In order to explore whether a participant's pre-interaction FS score was indicative of how participants physically engaged with PARO during their robot-interaction session simple linear regressions were calculated. All results were non-significant. Scatterplots, with lines of best fit, visualising this relationship between FS pre-interaction (where higher scores = higher feelings of felt security in participants (n = 60) immediately prior to interaction with PARO) and the three categories of touching behaviours, *no touch*, *active touch* and *intimate touch*, were produced (Figures 4.12, 4.13 and 4.14). Although non-significant, these plots show that participants' pre-interaction FS score trended positively with the proportion of normalised total time spent not touching PARO, or touching PARO in an active manner, whilst pre-interaction FS scores trended negatively with total time spent touching PARO in an intimate manner.

To explore whether there was a relationship between how a participant (n = 60) physically engaged with PARO during the interaction period and their FS delta score, simple linear regressions were calculated to predict FS delta score (dependent variable) based on each of the the three categories of touching behaviours, *no touch*, *active touch* and *intimate touch* (Table 4.7).

A significant regression equation was found on FS delta score of *no touch*,

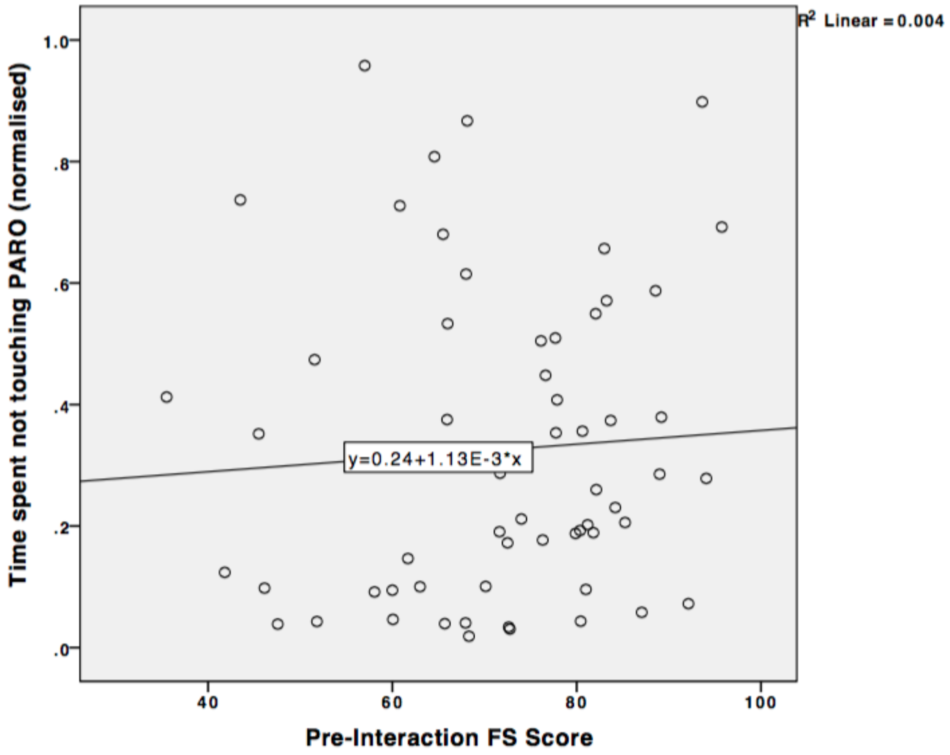


Figure 4.12. Scatterplot ($n = 60$) with line of best fit visualising the relationship between FS pre-interaction and touching behaviour coded as *no touch*.

$F(1,\ 58) = 18.27$, $p = .001$, with an R^2 of .24, and on FS delta score of *intimate touch*, $F(1,\ 58) = 16.37$, $p = .001$, with an R^2 of .22. These relationships are visualised in scatterplots (with lines of best fit) in Figures 4.15 and 4.16.

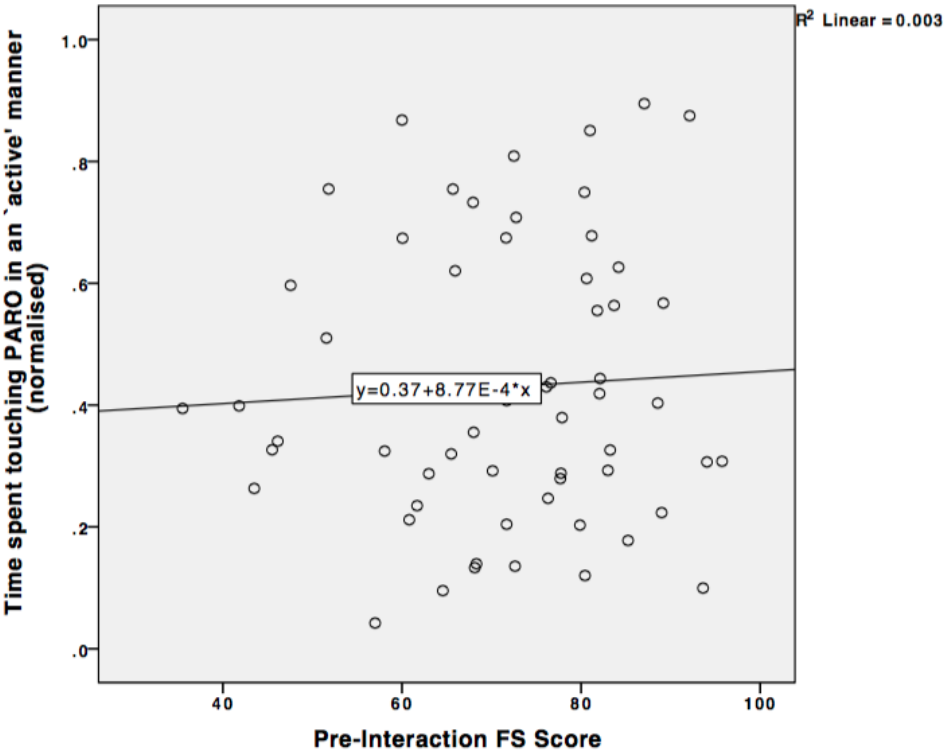


Figure 4.13. Scatterplot ($n = 60$) with line of best fit visualising the relationship between FS pre-interaction and touching behaviour coded as *active*.

Table 4.7. Simple linear regression ($n = 60$) to predict FS delta score (DV) based on each of the three categories of touching behaviours (normalised), *no touch*, *active touch* and *intimate touch* (IVs).

Touch Category	FS Delta Score F	FS Delta Score p
No touch	18.27	.001*
Active	0.11	.742
Intimate	16.37	.001*

[1] *Denotes significant effect.

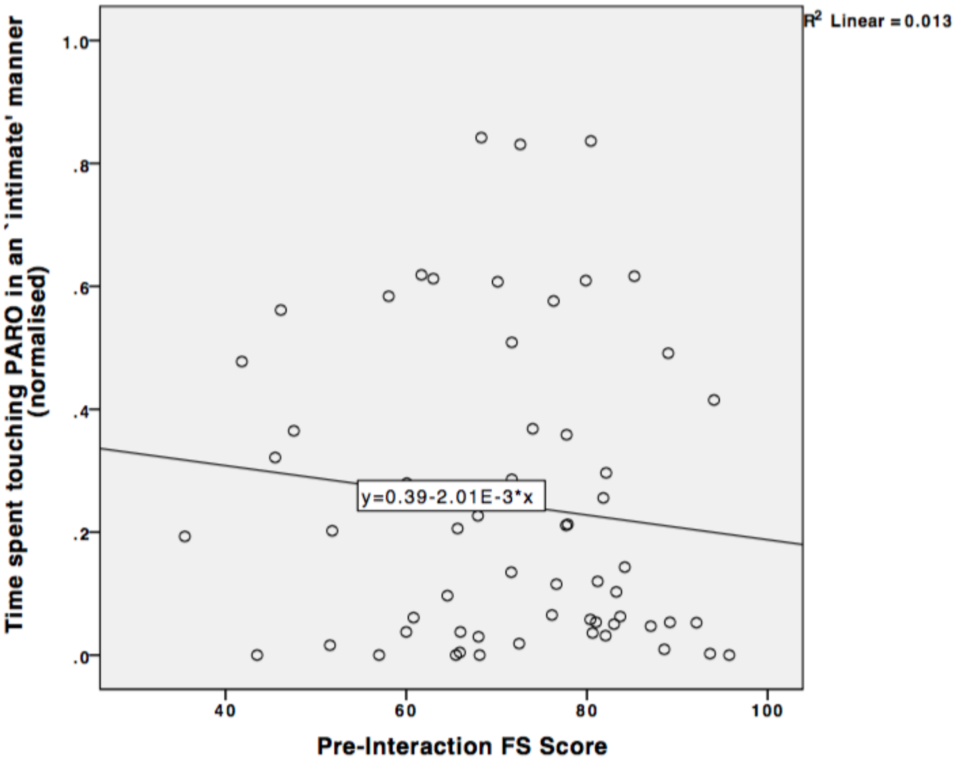


Figure 4.14. Scatterplot ($n = 60$) with line of best fit visualising the relationship between FS pre-interaction and touching behaviour coded as *intimate*.

Hypothesis 2: Relationships between physical engagement and FS scores are mediated by caregiving and attachment styles

To explore whether pre-interaction FS scores (dependent variable) were correlated with a participant's caregiving and attachment styles (independent variables), simple linear regressions ($n = 60$) were calculated (Table 4.8). Scatterplots of these interactions show all four are negative correlations (Figures 4.17, 4.18, 4.19 and 4.20).

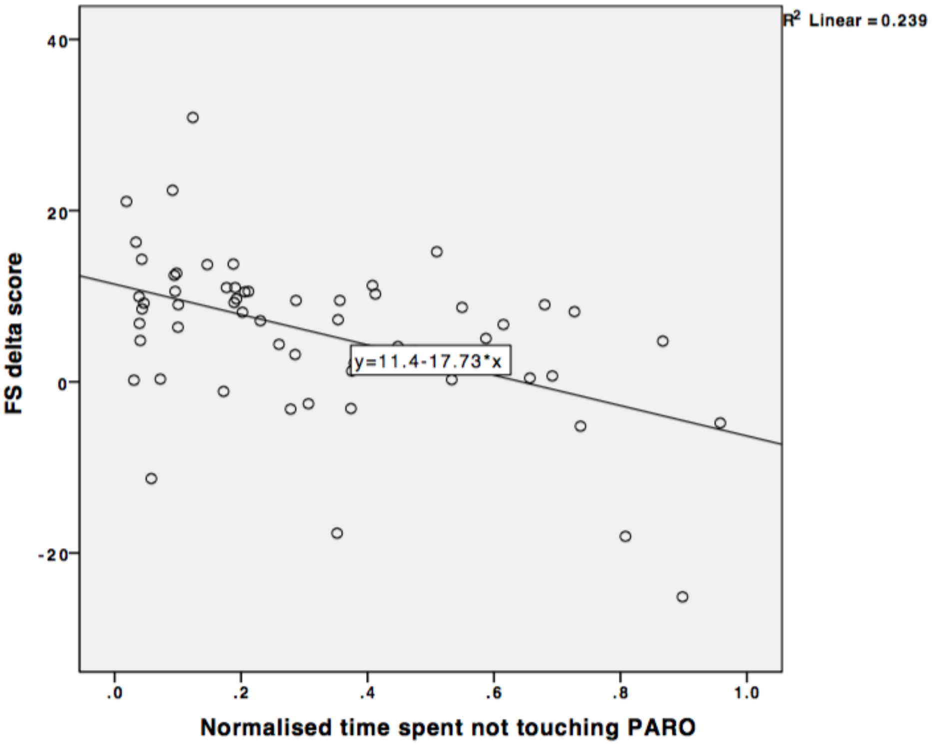


Figure 4.15. Scatterplot with line of best fit visualising the negative relationship between FS delta score ($n = 60$) and no touch engagement with PARO.

Simple linear regressions ($n = 60$) were also conducted to explore whether participant's FS delta scores (dependent variable) were correlated with a participants caregiving and attachment styles (independent variables). All results were non-significant (Table 4.9).

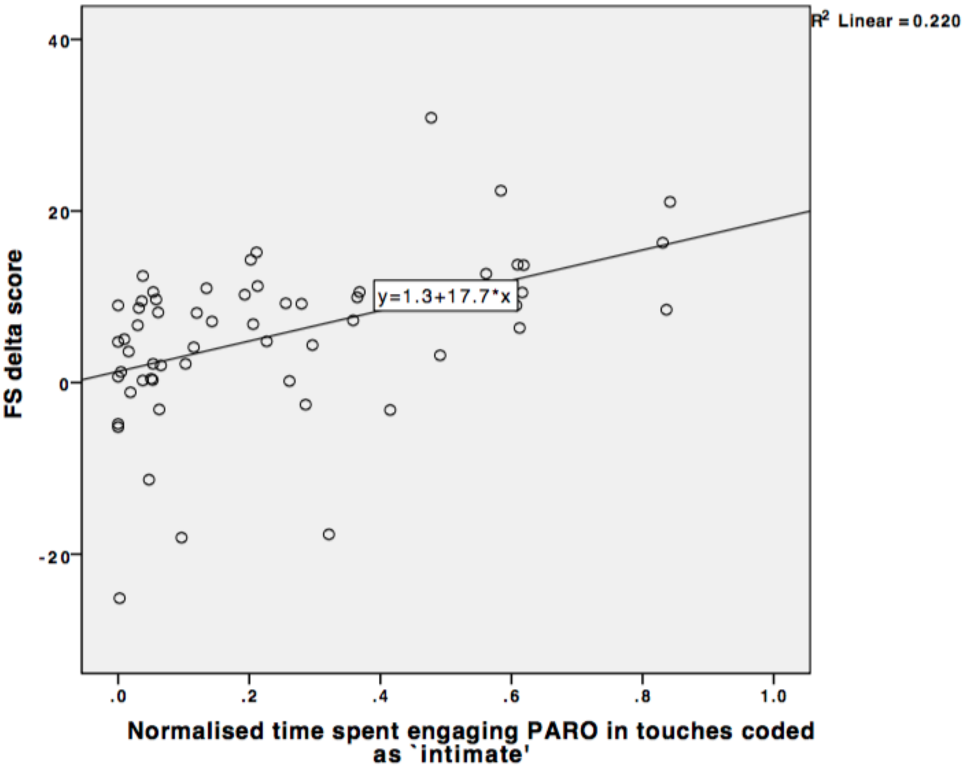


Figure 4.16. Scatterplot with line of best fit visualising the positive relationship between FS delta score ($n = 60$) and physical engagement with PARO coded as *intimate touch*.

Table 4.8. Simple linear regressions ($n = 60$) to correlate pre-interaction FS scores (DVs) against each of the four care-giving and attachment subscales (IVs).

Individual Differences Subscales	Pre-Interaction FS Score F	Pre-Interaction FS Score p
Avoidant ECR-S	2.03	.160
Anxious ECR-S	7.63	.008*
Deactivated CSS	1.80	.185
Hyperactivated CSS	4.32	.042*

[1] *Denotes significant effect.

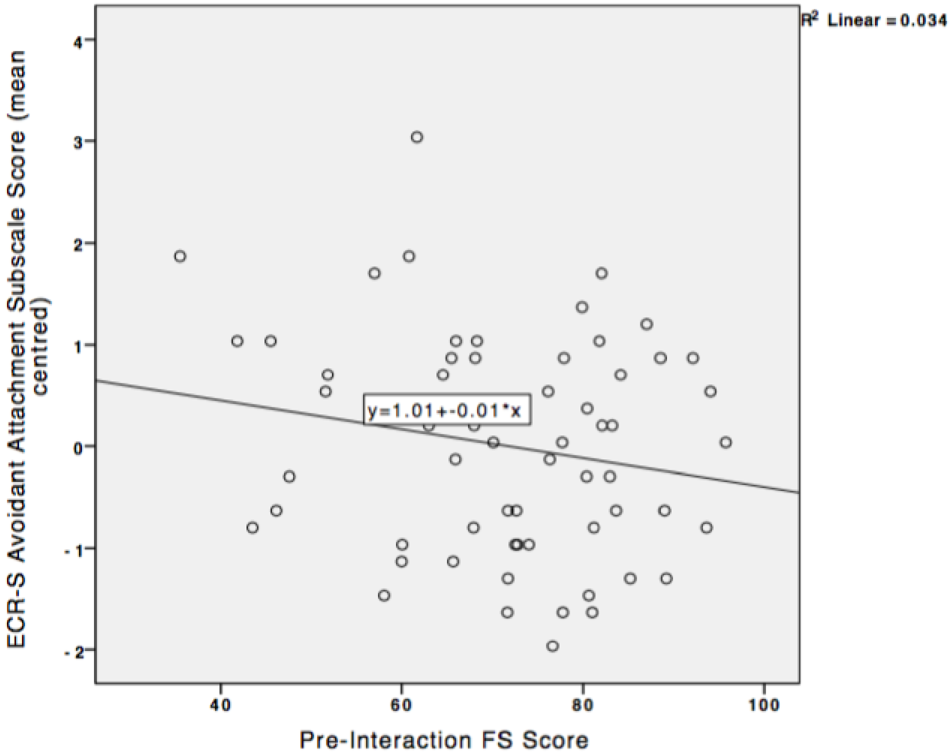


Figure 4.17. Scatterplot with line of best fit visualising the negative relationship between pre-interaction FS delta score (n = 60) and ECR-S Avoidant Attachment subscale scores.

Table 4.9. Simple linear regressions (n = 60) to correlate FS delta scores (DVs) against each of the four caregiving and attachment subscales (IVs).

Individual Differences Subscales	FS Delta Score F	FS Delta Score p
Avoidant ECR-S	0.00	.984
Anxious ECR-S	0.05	.818
Deactivated CSS	0.16	.688
Hyperactivated CSS	0.31	.577

In order to explore whether physical engagement with PARO as measured via the three categories of touching behaviour was mediated by the par-

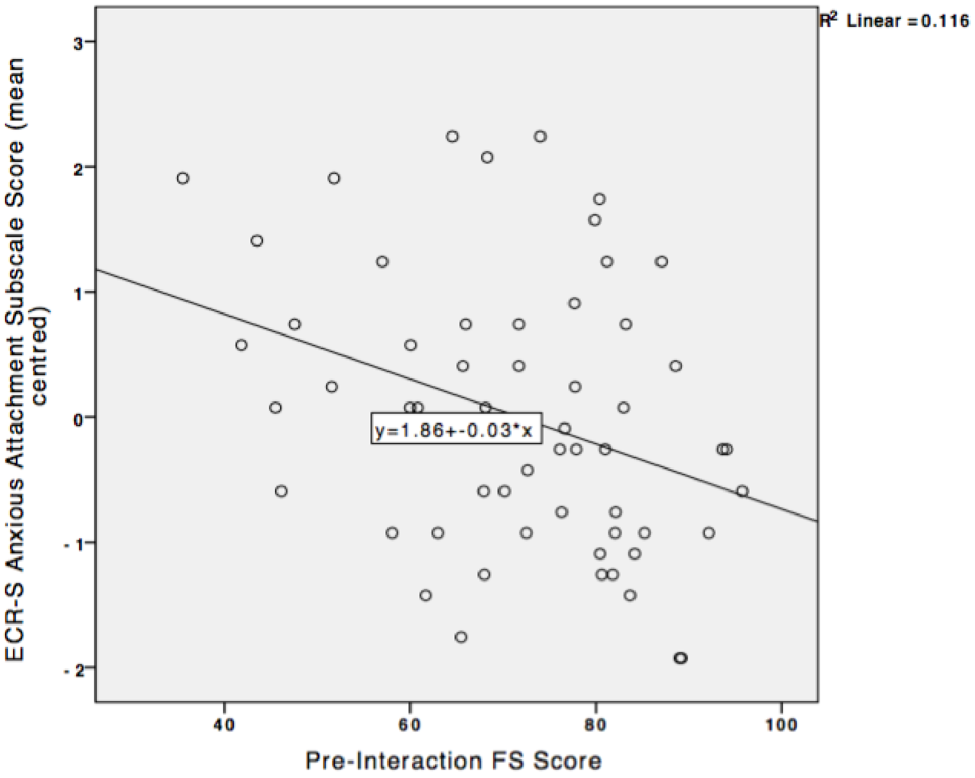


Figure 4.18. Scatterplot with line of best fit visualising the negative relationship between pre-interaction FS delta score (n = 60) and ECR-S Anxious Attachment subscale scores.

ticipant's caregiving and attachment styles, simple linear regressions (n = 60) were first calculated to predict amount of time (normalised) spent engaged with PARO in each of the three touch category behaviours (*no touch*, *active touch* and *intimate touch*: dependent variables) based on each subscale from the pre-session measures of individual differences (*ECR-S attachment avoidance; ECR-S attachment anxiety; CSS deactivated caregiving; and CSS hyperactivated caregiving*: independent variables). With two exceptions, all regression equations were non-significant (Table 4.10).

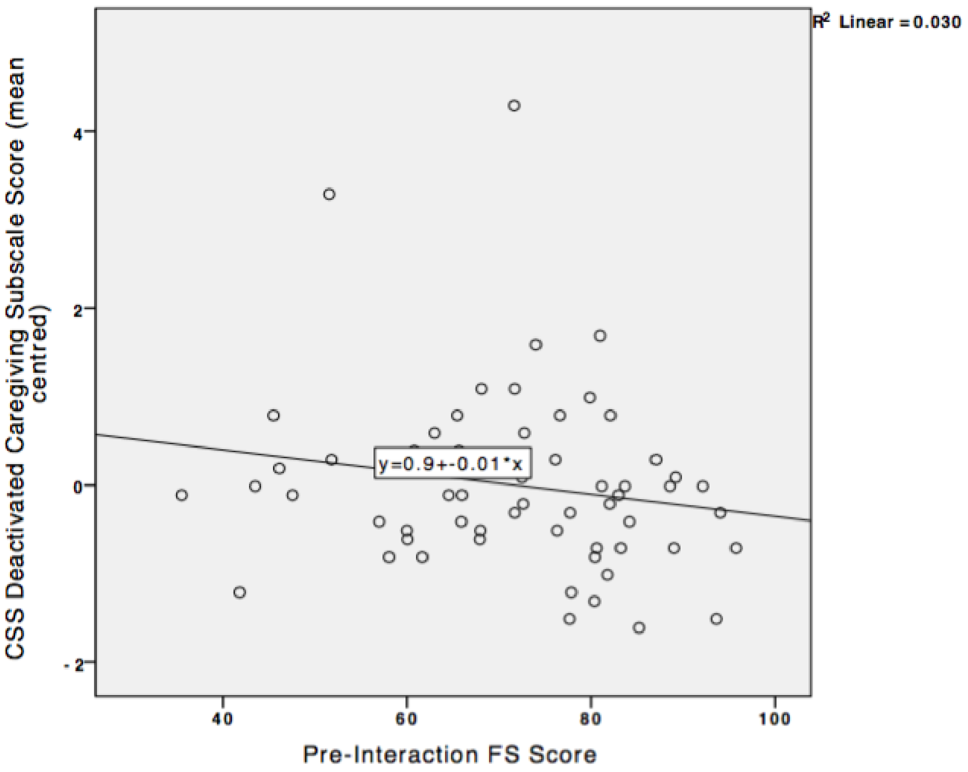


Figure 4.19. Scatterplot with line of best fit visualising the negative relationship between pre-interaction FS delta score ($n = 60$) and CSS Deactivated Caregiving subscale scores.

Table 4.10. Simple linear regressions ($n = 60$) to predict amount of time (normalised) spent engaged with PARO in each of the three touch category behaviours (DVs) based on each of the four individual differences subscales (IVs)

Individual Differences Subscales	NT *F*	NT *p*	Act *F*	Act *p*	Int *F*	Int *p*
Avoidant ECR-S	3.65	.061	5.14	.027*	0.01	.927
Anxious ECR-S	0.30	.589	0.02	.895	0.20	.660
Deactivated CSS	0.11	.739	2.47	.121	1.20	.279
Hyperactivated CSS	3.82	.055	5.74	.020*	0.02	.878

[1] ECR-S and CSS subscale values have been mean centred.

[2] *Denotes significant effect.

[2] NT = No Touch; Act = Active; Int = Intimate.

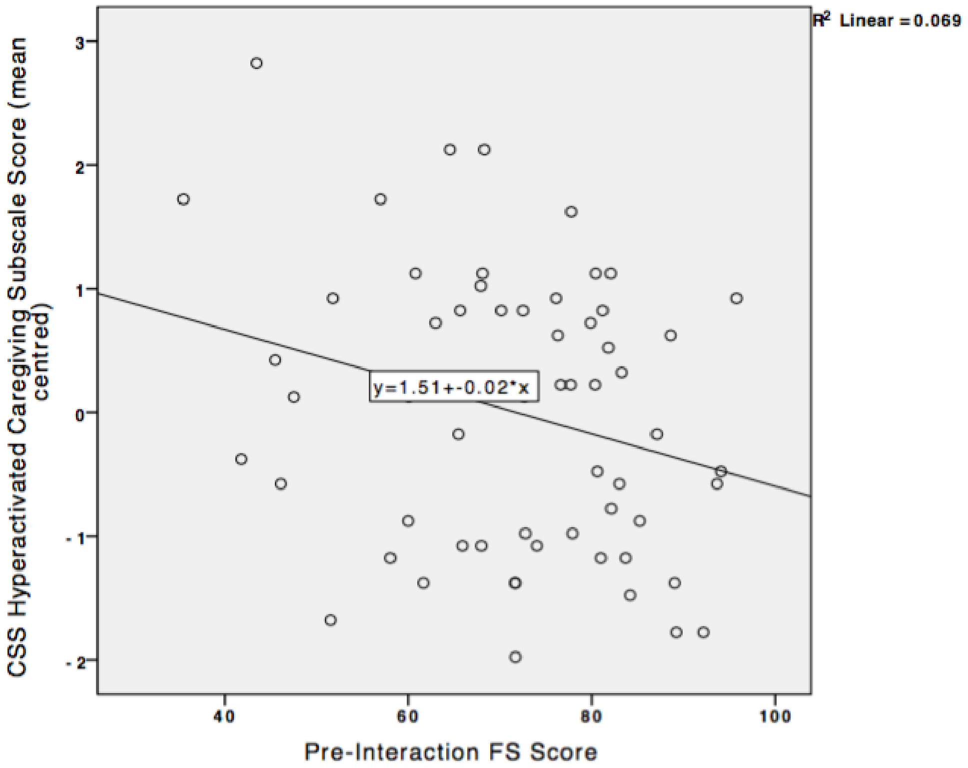


Figure 4.20. Scatterplot with line of best fit visualising the negative relationship between pre-interaction FS delta score ($n = 60$) and CSS Hyperactivated Caregiving subscale scores.

In order to run a mediation analysis several conditions must be met (Baron and Kenny, 1986). The independent (predictor) variables (which in the models being explored here would be either categories of physical engagement or pre-interaction FS scores) must be related to the dependent (outcome) variables (FS delta scores or categories of physical engagement) as well as the proposed mediator variable (which in both cases would be the individual differences subscales).[10] Additionally, the mediator variable (individual differences subscales) must be related to the dependent (outcome)

---

[10] Individual differences subscales include anxious and avoidant attachment scores from the ECR-S and deactivated and hyperactivated caregiving scores from the CSS

variables (either FS delta scores or categories of physical engagement). Although the *no touch* and *intimate touch* categories of physical engagement did significantly predict FS delta scores (see Table 4.7), there was no significant relationship between pre-interaction FS scores and categories of physical engagement (see Figures 4.12, 4.13 and 4.14). Further, the individual differences subscales were not significantly related to FS delta scores, and of 12 possible relationships between the subscales and the three categories of physical engagement only two were significant (see Table 4.10). These results may be due to lack of variation in individual differences subscale scores in this sample (see Figures 4.21 and 4.22, which visualise the spread of participant's CSS and ECR-S scores), nonetheless further exploration of the relationships between individual differences, physical engagement and felt security using mediation analysis was not appropriate.

Instead hierarchical multiple regression analyses were calculated to explore the impact of caregiving and attachment styles on three significant relationships: 1) the difference in pre- and post-interaction FS scores; 2) the positive relationship between FS delta and *intimate touch*, and; 3) the negative relationship between FS delta and *no touch*, with the caregiving and attachment style subscales included in each regression model as control variables.[11]

---

[11] The interaction terms of *ECR-S attachment avoidance* x *ECR-S attachment anxiety*, and *CSS deactivated caregiving* x *CSS hyperactivated caregiving*, as well as the IRI Fantasy disposition scores, were originally included in the hierarchical multiple regression analyses. These variables did not add any predictive power to the models and so were not included in the models here reported.

A multiple hierarchical regression was calculated to predict post-interaction FS scores based on pre-interaction FS scores whilst controlling for caregiving and attachment styles. A significant regression equation was found $F(5,\ 54) = 20.42$, $p <.001$, with an R^2 of .65, and a Durbin-Watson statistic of 2.343 indicating uncorrelated residual terms. The values reported in Table 4.11 reveal that post-interaction FS scores are significantly positively correlated with pre-interaction FS scores after controlling for all four of the caregiving and attachment style subscales.

A multiple hierarchical regression was calculated to predict FS delta scores based on *intimate touch* whilst controlling for caregiving and attachment styles. A significant regression equation was found $F(5,\ 54) = 3.20$, $p =$ .013, with an R^2 of .23, and a Durbin-Watson statistic of 2.318 indicating uncorrelated residual terms. The values reported in Table 4.12 reveal that FS delta scores are significantly positively correlated with the amount of time a participant spent physical engaging PARO in behaviours coded as *intimate touch* after controlling for all four of the caregiving and attachment style subscales.

A multiple hierarchical regression was calculated to predict FS delta scores based on *no touch* whilst controlling for caregiving and attachment styles. A significant regression equation was found $F(5,\ 54) = 3.76$, $p = .005$, with an R^2 of .26, and a Durbin-Watson statistic of 2.217 indicating uncorrelated residual terms. The values reported in Table 4.13 reveal that FS delta scores are significantly negatively correlated with the amount of time a participant spent physical engaging PARO in behaviours coded as *no touch* after controlling for all four of the caregiving and attachment style subscales.

Table 4.11. Output of the hierarchical multiple regression ($n = 60$) that related post-interaction FS scores (DV) to pre-interaction FS scores whilst controlling for individual differences subscales (mean centred).

	B	*SE B*	β
Step 1			
Constant	77.33	1.83	
ECR-S Avoidant Attachment	-1.81	1.72	-.13
ECR-S Anxious Attachment	-4.50	1.74	-.32*
Step 2			
Constant	77.33	1.80	
ECR-S Avoidant Attachment	-1.60	1.75	-.12
ECR-S Anxious Attachment	-2.87	1.94	-.21
CSS Deactivated Caregiving	-3.10	1.86	-.20
CSS Hyperactivated Caregiving	-2.66	1.89	-.20
Step 3			
Constant	20.37	6.80	
ECR-S Avoidant Attachment	-0.24	1.16	-.02
ECR-S Anxious Attachment	-0.45	1.31	-.03
CSS Deactivated Caregiving	-1.13	1.25	-.08
CSS Hyperactivated Caregiving	-1.17	1.26	-.09
Pre-Interaction FS	0.79	0.09	.75***

[1] $R^2 = .13$ for Step 1, *Delta* $R^2 = .06$ for Step 2, *Delta* $R^2 = .46$ for Step 3 (p <.001).

[2] Significance level of variable is denoted by asterisks, *p <.05, **p <.01, ***p <.001.

Table 4.12. Output of the hierarchical multiple regression ($n = 60$) that related FS delta scores (DV) to physical engagement coded as *intimate touch* (normalised) whilst controlling for individual differences subscales (mean centred).

	B	*SE B*	β
Step 1			
Constant	5.63	1.21	
ECR-S Avoidant Attachment	0.01	1.14	.00
ECR-S Anxious Attachment	-0.26	1.15	-.03
Step 2			
Constant	5.63	1.23	
ECR-S Avoidant Attachment	0.11	1.19	.01
ECR-S Anxious Attachment	0.17	1.33	.02
CSS Deactivated Caregiving	-0.62	1.27	-.07
CSS Hyperactivated Caregiving	-0.78	1.29	-.10
Step 3			
Constant	1.27	1.56	
ECR-S Avoidant Attachment	0.12	1.06	.01
ECR-S Anxious Attachment	0.37	1.18	.04
CSS Deactivated Caregiving	-0.05	1.14	-.01
CSS Hyperactivated Caregiving	-0.87	1.15	-.11
Intimate Touch	17.84	4.56	.47***

[1] $R^2 = .00$ for Step 1, *Delta* $R^2 = .01$ for Step 2, *Delta* $R^2 = .22$ for Step 3 ($p < .001$).

[2] Significance level of variable is denoted by asterisks, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 4.13. Output of the hierarchical multiple regression ($n = 60$) that related FS delta scores (DV) to physical engagement coded as *no touch* (normalised) whilst controlling for individual differences subscales (mean centred).

	B	*SE B*	β
Step 1			
Constant	5.63	1.21	
ECR-S Avoidant Attachment	0.01	1.14	.00
ECR-S Anxious Attachment	-0.26	1.15	-.03
Step 2			
Constant	5.63	1.23	
ECR-S Avoidant Attachment	0.11	1.19	.01
ECR-S Anxious Attachment	0.17	1.33	.02
CSS Deactivated Caregiving	-0.62	1.27	-.07
CSS Hyperactivated Caregiving	-0.78	1.29	-.10
Step 3			
Constant	11.81	1.81	
ECR-S Avoidant Attachment	0.97	1.06	.11
ECR-S Anxious Attachment	-0.09	1.16	-.01
CSS Deactivated Caregiving	-0.53	1.11	-.06
CSS Hyperactivated Caregiving	0.18	1.15	.02
No Touch	-19.00	4.47	-.52***

[1] R^2 = .00 for Step 1, *Delta* R^2 = .01 for Step 2, *Delta* R^2 = .25 for Step 3 (p <.001).

[2] Significance level of variable is denoted by asterisks, *p <.05, **p <.01, ***p <.001.

Finally scatterplots of CSS and ECR-S scores were created which visualised participants' ($n = 60$) caregiving and attachment styles, by mapping

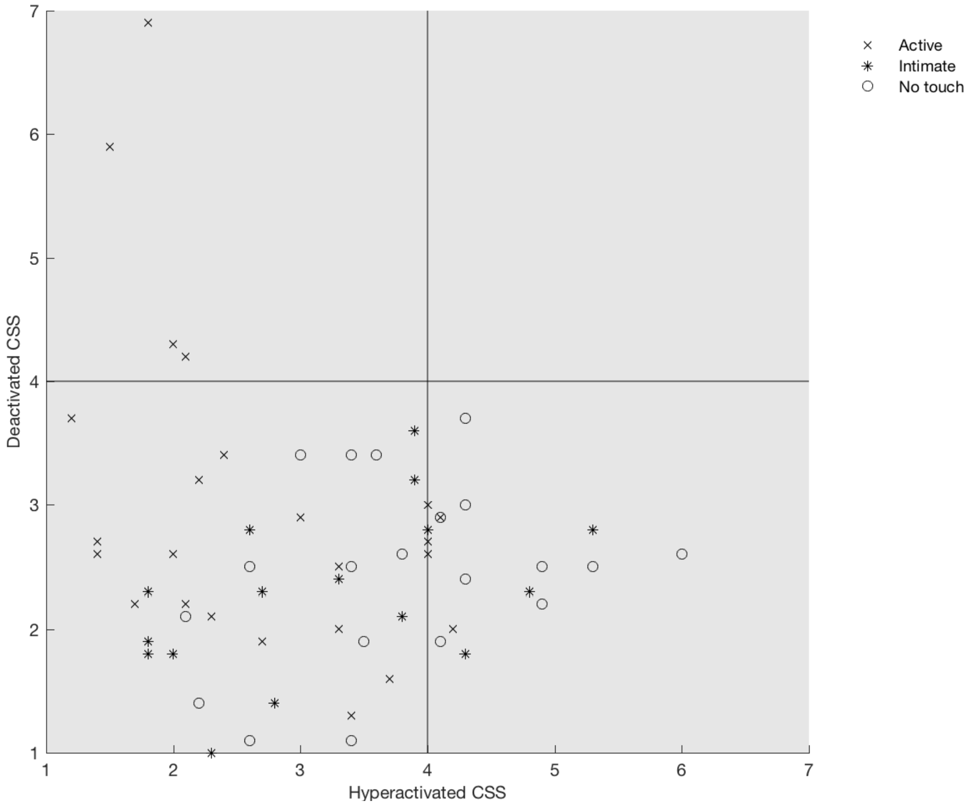


Figure 4.21. Scatterplot ($n = 60$) of relationship between CSS scores and three categories of physical behaviours coded for (*no touch*, *active touch* or *intimate touch*) within which participants spent the majority of their time whilst interacting with PARO. Note the two participants with overlapping scores but distinct physical engagement categories (*active* and *no touch*) in the anxious-hyperactivated caregiving quadrant (lower right).

their scores over the caregiving and attachment style quadrants which are created when deactivated or avoidant scores are combined with hyperactivated or anxious scores to create a total caregiving or attachment style score for each participant (see Figure 4.21 for CSS scores, and Figure 4.22 for ECR-S scores). Scatterplots display participants by whichever of the three categories of physical behaviours coded for (*no touch*, *active touch* or *intimate touch*) within which they spent the majority of their time whilst

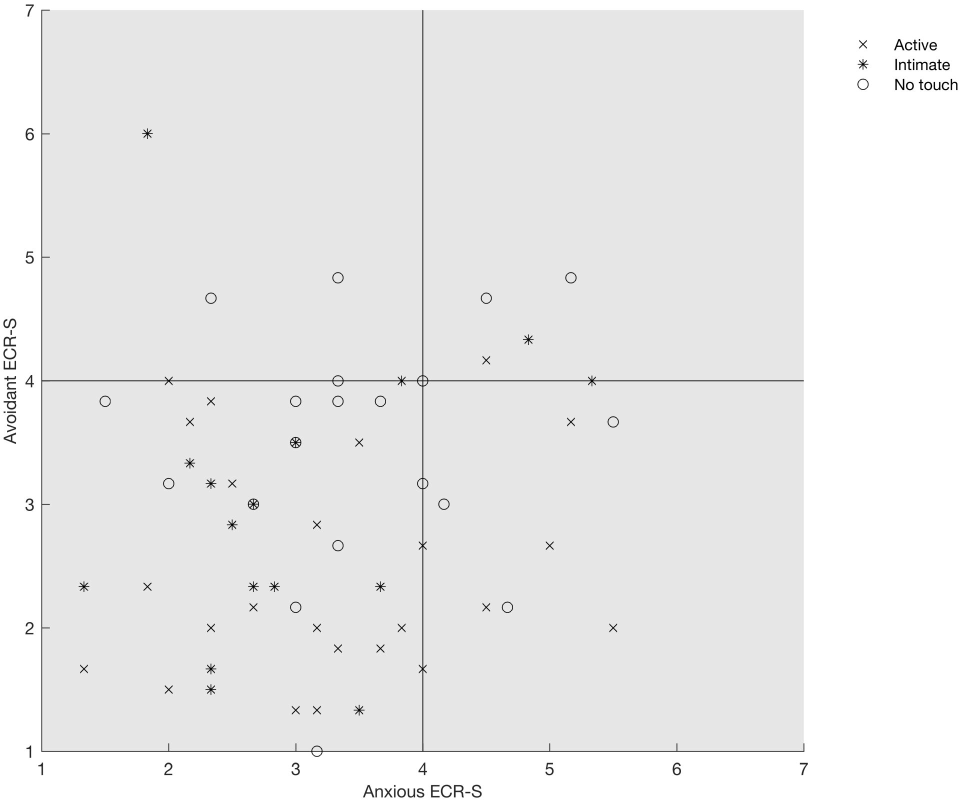


Figure 4.22. Scatterplot ($n = 60$) of relationship between ECR-S scores and three categories of physical behaviours coded for (*no touch*, *active touch* or *intimate touch*) within which participants spent the majority of their time whilst interacting with PARO. Note the two sets of participants with overlapping scores but distinct physical engagement categories (*intimate* and *no touch*) in the secure attachment quadrant (lower left).

interacting with PARO. Thus, for example, an individual who spent 30% of their total time with PARO not touching the robot, 30% actively touching the robot, and 40% intimately touching the robot would be categorised as an *intimate touch* participant and positioned on the scatterplot at the location of their overall caregiving or attachment subscale style.

Note in Figures 4.21 and 4.22 that as in the pilot, the majority of participants in the main study fall in the optimal caregiving/secure attachment style quadrants (bottom left), and area which indicates that a participant had low scores in each dimension.

4.6.3 Discussion

The primary aim of this study was to explore whether an interaction period with the social biomimetic robot PARO would lead to a positive change in a participant's felt security, and as with the pilot study, this main study demonstrated that an interaction period with a PARO does correlate with positive increases in a participant's FS score from pre- to post-interaction.

This study also sought to investigate why PARO might have such an impact on FS via two hypotheses.

The first hypothesis was in two parts, and addressed the potential impact of different types of physical engagement with the robot upon FS. First it was hypothesised that participants with lower pre-interaction FS scores would physically engage PARO less during the interaction period than participants with higher pre-interaction FS scores. Scatterplots visualising the relationship between FS pre-interaction and each of the three categories of touching behaviours observed during the interaction period: *no touch, active touch* and *intimate touch*, showed no strong correlations between type and amount of behaviour engaged in and FS pre-interaction scores (see Figures 4.12, 4.13 and 4.14). Secondly it was hypothesised that participants who spent the majority of their interaction time physically engaging PARO via more proximate behaviours (such a caressive strokes, coded as *intimate touch*) would have higher FS delta scores than individuals who either did not engage PARO (*no touch*) or spent their time with PARO exploring the robot (*active touch*). Here, the study revealed a significant relationship between individuals who spent the majority of their time not touching PARO and low FS delta scores, and individuals who spent the majority of their time touching PARO in an intimate way, by for example stroking the robot, and high FS delta scores (see Figures 4.15 and 4.16). One issue with this finding is that although there was no significant relationship between how a participant engaged PARO and their FS pre-interaction score, there was an observable negative correlation between time spent engaging PARO intimately and FS pre-interaction scores, meaning that individuals who engaged PARO intimately were more likely to have had lower FS scores to begin with, and would therefore be able to attain a greater FS delta score than an individual whose pre-interaction (that is, baseline) FS score was higher. However, given the lack of significance with the first half of this hypothesis (there is no relationship between pre-interaction FS and touching behaviours observed) this may not be relevant to the subsequent significant result.

The second hypothesis was that physical engagement with the robot as measured via touching behaviours would be mediated by the participant's caregiving and attachment styles. Such that, participants who scored more highly on deactivated or avoidant styles would spend the majority of their interaction time either not touching PARO or engaging PARO distally, for example by exploring the robot instead of caressing it, and that those participants would then have smaller FS delta scores. Whilst participants who scored more highly on hyperactived or anxious styles would spend the majority of their interaction time engaging PARO caressively, and in turn would have larger FS delta scores. However, in order to run a mediation analysis several conditions must be met (Baron and Kenny, 1986). As reported in the results section above, these conditions were not met in this study. Instead hierarchical multiple regression analyses were calculated to explore the impact of caregiving and attachment styles on the difference in pre- and post-interaction FS scores, the positive relationship between FS delta and *intimate touch*, and the negative relationship between FS delta and *no touch*. This batch of analyses revealed that even after controlling for individual differences in caregiving and attachment style, time spent with PARO is significantly related to positive increases in FS, whilst more time spent engaging PARO in intimate touches leads to greater FS delta scores, whilst spending more time not touching PARO at all leads to smaller FS delta scores. Thus, not only did engaging with PARO positively effect felt security, but how a person engaged with PARO also impacted their outcome felt security level, whilst caregiving and attachment style could not be used to predict how a person would engage PARO, and thus neither could it be used to predict how much engaging with PARO might impact their post-interaction felt security.

It should be noted that the individual difference's measures were accurately capturing what they were designed to, such that there are relevant correlations (non-significant) between attachment and caregiving scores, and pre-interaction FS scores. Both anxious ECR-S scores and hyperactived CSS scores were significantly negatively correlated with pre-interaction FS scores, and avoidant ECR-S scores and deactivated CSS scores were also negatively correlated, though not significantly (see Table 4.8, and Figures 4.17, 4.18, 4.19 and 4.20). These correlations are expected, as the Felt Security Scale was designed to capture the full dimensionality of felt security, which is inclusive of attachment security (Gillath et al., 2009). Attachment security measures fluctuations in adult attachment style conceptualised as falling along dimensions of anxiety, avoidance and security. Individuals scoring more highly in anxiety or avoidance (and by extension hyperactivated and deactivated caregiving) would be expected to have lower security scores. Hence scores from the FSS should negatively correlate with those from the ECR-S and the CSS.

The study shows a broad variation in how individuals physically engage PARO, as can be seen in the scatterplots visualising participants by CSS (Figure 4.21) and ECR-S (Figure 4.22) scores and the behaviour they spent the majority of their interaction time engaging in. There is also an overall improvement in FS to be seen after an interaction time with PARO. However, why individuals engage with PARO as they do, such that their outcome FS is impacted, is not fully captured by the analysis of individual differences in this study. It should further be noted that there were no differences between males and females in how they engaged PARO and their corresponding FS delta scores.

Further considerations

The findings from this study warrant further consideration. Regarding the experimental set-up there are three clear limitations, which were the study to be conducted again should be addressed. The study could have benefitted from two control conditions, wherein two other groups of participants were exposed to an interaction period with a PARO that was off, and another interactive activity such as a game, in order to confirm that it was the active PARO which was contributing to the positive FS score increases, and to further compare those changes in FS delta scores to another engaging activity. The ECR-S questionnaire used in the pre-interaction online section of the study could have used global style questions instead, replacing the term ‘romantic partner’ with alternatives such as ‘others’ or ‘those close to me’. Questioning participants about their feelings towards ‘romantic partners’ without confirming whether or not they had experience of such relationships was an oversight. However, as discussed by Luke et al. (2012) although there is a link between people who suffer from social anxiety, and who may therefore not have as much experience of close relationships as those who do not suffer from anxiety, and anxious and/or avoidant attachment styles (Darcy et al., 2005), past research has demonstrated that priming secure relationships increases feelings of felt security across participants nonetheless (Rushforth, 2009), which may point to the flexibility in participant's ability to interpret and answer the ECR-C regardless of the pronouns used. Finally, it is worth noting that the sample used in this study came from a self-selecting group of individuals who most likely already had a preference for liking robots. This invites criticism of a bias in the sample, as participants may have been happy to interact with any robot, and felt more secure in themselves after the interaction regardless as they were nervous, and excited, to meet an unknown robot pre-interaction.

Analysis of the video data revealed a large variation in how individuals were interacting with PARO, which the coding scheme was an attempt to reflect. However, the reason behind such variation was not captured by the individual difference's measures taken in the study. One reason for his could be the lack of variation in the sample itself. As can be seen in the scatterplots in Figures 4.21 and 4.22, the majority of participants fell into the securely attached and optimal caregiving quadrants. Perhaps no link between attachment and caregiving styles and physical engagement measures was found because the sample was not from a non-secure group. Future studies within populations such as those suffering from clinical anxiety problems may result in different conclusions.

Large variation in observable behaviour with the robot may also be due to the novelty of robots themselves. Unlike other people, or even animals, the cognitive model of robots held by individuals who do not have experience with robotics, as the sample in this study were, may not be consistent. Thus, typical behaviours displayed by individuals interacting socially with known agents may well differ to those behaviours displayed whilst interacting with an unknown social agent, such as a social biomimetic robot. People create cognitive models of their environment over time, which repeated exposure weakens (Summerfield et al., 2008). As individuals learn to predict their environmental input they become less sensitive to said environments, but novelty upends this template. Biomimetic robots can be described as exhibiting properties which straddle the boundaries between animacy and inanimacy, as demonstrated by Jipson et al. (2016), in their study exploring the ontological status as a robotic dog, which was talked about by adults and children using both animate and artefact properties. Findings such as these could lead to the conclusion that the novelty of a social robot, a liminal agent, may produce unpredictable behaviours from the human-half of the observed human-robot dyad because such robots are not built into the predictable models of the world that individuals have.

However, it may not be the robotic novelty of PARO that produces such a wide variety of observed physical behaviours from individuals. As discussed in Chapter 2 (Section 2.4.2) the pairing of a client with the appropriate AAT animal is a nuanced process, which takes client temperament into consideration on a case-by-case basis. Although attachment style may be considered when pairing an individual with an animal, based on what is known about how individuals with different attachment styles may respond to social agents, the final decision may be left up to what is observed during the meeting between client and animal itself. Knowledge of a client's attachment style, though informative for the development of a goodness of fit between a client's individual attachment needs and their relationship with their AAT animal, remains advisory and is not prescriptive (see Chapter 3, Section 3.3.2). To illustrate this point take for example a situation in which two people are observed interacting with a potential AAT dog on two separate occasions. Person A greets the dog exuberantly, and goes on to display a number of behaviours which this study would code as *intimate touch* physical engagements, such as stroking the dog playfully. Person B engages the dog in exactly the same manner. Person A scored highly on the anxiety dimension of the ECR-S, whilst Person B scored highly on the avoidant dimension. After discussion with their therapist it is concluded that the exuberant relationship with the dog is a good fit for both clients. The therapist, in discussing the observed behaviours with each client, has learnt that Person A being an anxious individual enjoys playful tactile stimulation, and reflectively also displays clinging behaviour with their romantic partner. Person B however, being a more avoidant individual, has a tendency to inhibit their proximity seeking behaviours with other humans having experienced traumatic relationships in the past, and as such seeks comfort in being able to express proximal behaviours with the dog, which represents a non-threatening and non-judgemental social agent. Different people may respond to the same animal in the same way, producing behaviours which are objectively identical, but that does not mean that the intentions behind those behaviours are the same.

Although past research has demonstrated that there is a large range of individual preferences exhibited by participants giving their opinion on zoomorphic robots regardless of, for example, participant gender or pet-ownership history (Jones et al., 2008), it can be posited that biomimetic robots may be perceived of as more animal-like than object-like (as discussed in Chapter 3). Thus, as in the example above it may be that participants in this study were responding to PARO as though it were an animal, and displaying appropriate nuanced behaviours in turn. Behaviours which fell outside simple categorisation of anxious or avoidant, proximity or inhibited proximity seeking.

4.7 Conclusion

The work in this chapter was an opportunity to explore a new methodology for measuring physical engagement with a social biomimetic robot, based on comparative understandings in HHI and HRI. Although the research findings were not conclusive regarding the involvement of individual differences in predicting behaviours observed during a human-PARO interaction, this study did demonstrate the potentially wide reaching application of robots for assistive therapy. The positive correlation shown between an intimate interaction with PARO and an increase in an individual's feelings of felt security occurred regardless of individual differences in attachment and caregiving style, and the participant's gender. If the PARO therapeutic robot can have an equally positive effect across a range of individuals, it can be seen as a therapeutic agent whose benefits are not restricted to only particular subgroups. In sum, felt security is increased after an interaction with PARO, and intimate physical engagement, as opposed to no engagement at all conducted during the interaction appears to be important to this positive effect.

4.7.1 Implications for subsequent chapters

The exploratory work in Chapter 4 was wide reaching, asking what features of users might be contributing to the positive increases in their felt security after an interaction with a biomimetic robot. For the next chapter the focus will shift to a discussion of what mechanisms of effect from the robot might be contributing to positive increases in user's felt security. Given the importance of engagement with a robot highlighted by Chapter 4, the thesis will now focus on what features of a robot might be having an impact on user felt security via increasing engagement with its user. Does it matter how predictable or believable the robot is? Humans can feel threatened by unpredictable behaviour, does that hold with robot behaviour? In order to explore precise mechanisms such as behaviour with a robot intended for RAT, Chapter 5 will discuss the importance of using a platform whose mechanisms can be selectively modified when conducting HRI studies. Chapter 5 will introduce the social biomimetic robot MIRO, whose open access control architecture is suitable for such systematic investigation. This is as opposed to PARO, whose control architecture does not allow for systematic manipulation, making the platform a poor choice for use in controlled experiments focussed on the mechanisms of the robot itself.

Chapter 5

MIRO: A Biomimetic Platform for Systematic Investigation of RAT

5.1 Summary

Chapter 4 demonstrated one way in which RAT is effective at leading to a positive psychological change in a user. PARO, the platform used in Chapter 4, is a closed platform that displays only simple behaviours. For an experiment which aims to discover mechanisms of effect researchers require a robot that can be selectively modified in order to test for features that are most effective for a particular use case, such as RAT. Therefore, in Chapter 5 the selectively controllable biomimetic platform, MIRO, is introduced. The chapter also presents an experiment validating one aspect of MIRO's design: its ability to communicate affect via light. The chapter highlights the platform's suitability for conducting controlled HRI studies. In discussing this the chapter also makes suggestions for the experimental work that will be presented in Chapter 7: a cross-cultural exploration of the importance of predictable mammalian-like behaviour for the promotion of emotional engagement.

5.2 Introduction

Chapter 4 demonstrated that RAT conducted with a biomimetic robot can positively impact a user's well-being. The robot used, PARO, displays only simple behaviours, and is a closed platform cannot be systematically manipulated (see Chapter 4, Section 4.3). However, in order to conduct controlled experiments, which explore precise mechanisms of effect, a robot whose features are fully accessible is required. Thus, in order to follow on from the work reported in Chapter 4 and explore what features of a social biomimetic robot have a positive impact on user felt security, it was necessary to acquire a robot that can be selectively modified.

As the follow up work for Chapter 4 was being conceived a new robot platform, MIRO, was being developed at Sheffield Robotics with external partners. MIRO's control system is a brain-based model with a layered architecture (Prescott et al., 1999). This results in a platform whose gross behaviour emerges from the competition between various sub-systems to explore locations in its environment with high sensory salience, to escape from stimuli that are perceived as 'threatening', to seek out goals, and to have highly responsive interactions with human users. The MIRO system is fully accessible, allowing for selective modifications to its behavioural features. This makes it an ideal platform for conducting the sorts of controlled experiments needed to explore mechanisms of effect in robots intended for RAT.

Part of the problem facing AAT researchers is the difficulty in knocking out single features of animals, whilst maintaining an otherwise fully functioning AAT compatible animal, in order to test the feature's impact on different scenarios. Using robots such as MIRO, whose brain-based control system lends itself to selective manipulation in order to conduct constrained experiments, is an important step in resolving this gap in the field, and allows therapeutic experiments to be conducted on single features whilst maintaining an otherwise identical unit (which as stated, is something that cannot be done with an AAT co-therapist). As discussed in Chapter 2, in a similar manner to the lack of controlled AAT studies, there is a lack of HRI literature which explores specific features of robots used in therapeutic situations. Further, HRI studies often suffer from a lack of specificity when it comes to the performance measures used (see Chapter 3), especially with respect to evaluating the role of design, and there is little formative evaluation in HRI due to a lack of tested methods for conducting rigorous and repeatable experiments (e.g., see Baxter et al. (2016); Clarkson and Arkin (2007); Riek (2012)). One general problem underlying these issues is that researchers vary too many parameters in a single experimental condition.

A good example of this common HRI issue of varying too many parameters at once is Jones et al. (2008), who reported a study that attempts to explore the suitability of animal-like features in a zoomorphic robot (in this case a modified iRobot Roomba, see Figure 5.1). The study is focussed on, "zoomorphic appearances and dog-like behavioural properties," of a modified robot unit, and the impact these might have on a participant's assessment of the robot's ability to locomote at their command towards a letter placed on the floor.

Figure 5.1. Modified iRobot Roomba as used in Jones et al. (2008). Unit shown is that for Condition 1: canid mechanisms included a tail that could be raised, lowered and wagged; ears that covered wireless speakers which emitted barking and whining sounds; and eyes that moved. A spotty fur covering was used to create a zoomorphic appearance.

These zoomorphic and dog-like properties (a tail that could be raised, lowered and wagged; ears that covered wireless speakers which emitted barking and whining sounds; eyes that moved; and a spotty fur covering used to create a zoomorphic appearance) were grouped as a single variable in a single condition. This was compared against Condition 2, in which the Roomba, though furry, did not vocalise nor animate its modifications; and Condition 3, an unmodified Roomba. The aim of the research was to establish whether participants would rate the performance of the robot more sympathetically if it were more ‘dog-like’.

There are two clear issues with this study. The first is that Condition 1 includes more than one modification. Thus any conclusion made from the work must make reference to all modifications, both audial and visual, and their co-existence. The ‘dog-like’ property of the robot in this study is therefore restrained to a very specific set of variables which cannot be treated independently. This is problematic for researchers who might want to use this work to establish design parameters for a zoomorphic robot, as it begs the question of which of the modifications that differ between the conditions are relevant to the performance assessments given by participants. For some users the sounds the robot emitted might have been influential, for others the movement of the modifications, and then the movements of what in particular? This leads to the second issue with the study: how the robot’s performance was assessed. The participants gave self-reported performance assessments based on how well the robot fetched the letters, and how frustrated and excited they were with the robot’s performance. The difficulty in interpreting these results, and the extent to which answers to these questions can be generalised beyond the very specific situation in which they are given, is limited. This is due to the subjective nature of self-reporting, and also the lack of specificity as to what exactly is being assessed. For example, if the robot was for some reason based on its visual or audial presentation perceived sympathetically it might then follow that in spite of it getting the task wrong it would be more favourably rated than an unmodified robot which might be treated less sympathetically for making mistakes. This is a helpful comparison, if the reason for the modified robot's sympathetic reception can be clearly ascertained. This is not possible if there is more than one variable change between conditions. Indeed, in the results there was a greater distribution of participants' ratings of the performance of the robot in Condition 1 when it performed the fetching task well, indicating perhaps a confusion on behalf of participants who were unsure what performance exactly they were meant to be rating: how well a robot locomoted to the correct letter, or how animal-like it was behaving, or a combination of the two? Contrast that with the same questions being asked of participants assessing the performance of a Condition 3 unit: an unmodified Roomba. Would participants be rating this robot based on how well it locomoted, and if so in what way would they rate the unit's performance of the task knowing that this exercise is not the function of a Roomba. Or were participants rating it on how 'robot-like' it was whilst performing its task, in which case if it failed at its task would participants then rate it unsympathetically, as the purpose of a mechanical robot would, presumably, be to attend to its task in a precise manner and nothing more. These questions highlight the difficulty in measuring the performance of a robot. Given the confusion that these questions raise as to what is really being measured, it is perhaps not then surprising that the work by Jones and colleagues concluded that the appearance of the robot over the three conditions had no significant effect on the participants' rating of the robot's performance.

Due to the proliferation of HRI experiments which attempt to vary too many parameters at once, such as this one by Jones and colleagues, the question remains, what aspects of design are having an impact in RAT? Heerink et al. (2013) conducted exploratory research in order to establish which features are needed for a robot intended for RAT to be successful at, for example, improving a user's sense of well-being. Their work produced a list of features compiled after interviewing 36 professional caregivers, both experienced or not with RAT. Required features included the animal-like unit having pettable fur, having realistic morphology and movement, looking innocent, being autonomous and adaptable, and being easy to use. However, even after comprehensively researching these features with users the authors concluded that much more research needs to be done on the listed requirements to focus on their specification.

One feature which this thesis has drawn attention to is emotional engagement (EE). In Chapter 4 individuals who touched PARO in a manner coded as *intimate* had greater felt security delta scores than individuals who spent the majority of their time interacting with PARO by not physically engaging it at all. This begs the question as to what design aspects in particular promote EE, such that a positive psychological change is experienced as a result? As explained above, it is difficult to selectively manipulate individual features of PARO in order to explore which ones are important for EE promotion. For example, that PARO has pettable fur might be important for encouraging a user to physically engage with it. The research by Heerink et al. (2013) would support that. However, it should be noted that the work by Shiloh et al. (2003) suggested that in an AAT animal fur may not be a primary factor in anxiety and arousal reduction (see Chapter 2, Section 2.4.1). It might then be because PARO is biomimetic in form. PARO looks alive, but it does not resemble an animal with which many users may have experience of (namely seals). Given this, users may not have an elevated expectation of the behaviour of the robot, which could otherwise lead to user frustration when the robot in turn fails to behave as expected (see MacDorman and Ishiguro (2006) for a discussion of the importance of visual and audial realism of robot acceptance; but also Jones et al. (2008) for a discussion of the disturbing effects on users caused by robot's displaying incongruous behaviours).

Another feature which may promote EE in a user is the predictability of the robot's behaviour. Humans appreciate predicability in animal behaviour, and stable behaviour is a requirement of a certified AAT co-therapist (Connor and Miller, 2000). Indeed, this question of the importance of predictable mammalian-like behaviours from a social biomimetic robot will be the focus of Chapter 7. Further, there is still much to be pursued with the question of how the individual differences of the users themselves impact how they interact with the robot. It is not only design aspects that promote EE but also a willingness to engage on behalf of the user. One aspect driving individual differences in response to a robot may be the social culture in which an individual has been brought up and lives in, influencing their expectations of robots, and shaping their mental models of machines. This too will be discussed in Chapter 7.

For the remainder of this chapter the biomimetic robot MIRO will be introduced, and an experiment conducted to validate one aspect of its design will be presented.

5.3 The Robotic Platform: MIRO

The MIRO robot was commissioned as a commercial pedagogical and leisure product, targeted particularly at the domestic and school markets.[1] MIRO is also intended as an artefact to drive public engagement with science, robotics in particular, and biomimetic robotics most of all.

5.3.1 Aesthetics and morphology

MIRO's aesthetics and morphology (Figure 5.2) were chosen to be engaging through evocation of a mammalian identity. Design choices explicitly avoided targeting a particular mammal in a bid to lower user expectation of behaviour and performance by comparison to a direct biological correlate. However, visual and behavioural cues were taken from puppies, kittens and rabbits. MIRO's personality aesthetics were required to capture the essence of the Japanese word *kawaii* ('cute'), whilst not being too

---

[1] The MIRO project was initially funded by Eaglemoss Publishing, and developed by a partnership with Eaglemoss, Sheffield Robotics, Sebastian Conran Associates, Gadgetlab and Buzzamo. MIRO remains a flagship Sheffield Robotics biomimetic development platform.

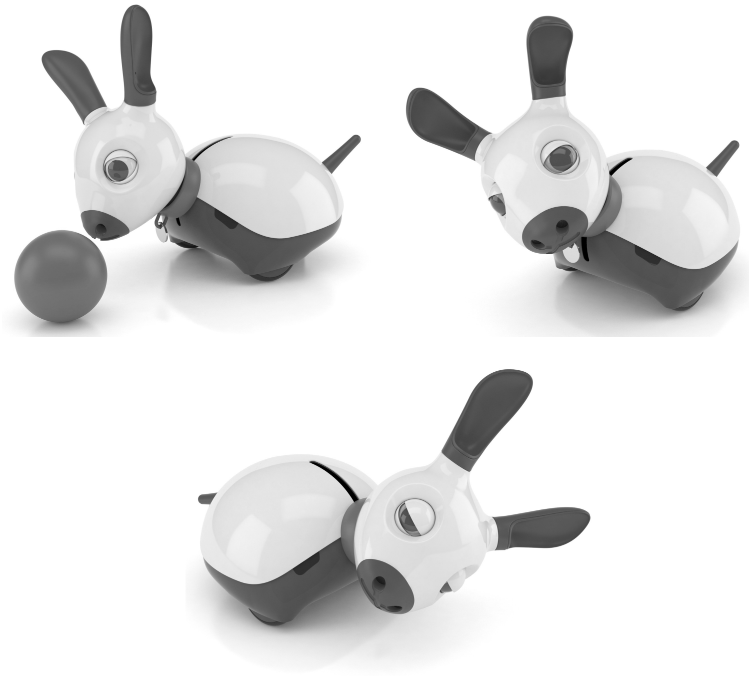


Figure 5.2. Concept art for MIRO expression of emotion through biomimetic, in particular mammalian-like, body language (imagery from Sebastian Conran Associates, Kensington, London, UK).

toy-like. The end result is intended to be somewhat of a ‘generic mammal’, though some specificity is naturally unavoidable (anecdotally, researchers at Sheffield Robotics have found that individuals interacting with MIRO consider the unit to be an amalgamation of many different mammals, from cows and donkeys to dogs and rabbits; few individuals liken MIRO to any one particular animal).

The platform is equipped with some of the same expressive appendages available to many mammals allowing mammal-like direct signalling of emotional state and responses to stimuli, for example blinking eyes, articulating neck, wagging tail and ears that move in a distinctly animal-like manner. MIRO also possesses a coloured lighting communication channel on its body and head that has no mammalian correlate. This channel also displays emotional responses to human interaction such as: happiness, enthusiasm, and sulking.

5.3.2 Platform

The MIRO platform is built around a core of a differential drive (plus caster) base and a three-DOF (lift, pitch, yaw) neck. Additional DOFs include two for each ear (curl, rotate), two for the tail (droop, wag), one for the caster (raise/lower), and one for the eyelids (open/close). Whilst these latter DOFs target only communication, the movements of the neck and body that serve locomotion and active sensing play a significant role in communication as well. Finally, the platform is equipped for sound production.

All DOFs are equipped with proprioceptive sensors (potentiometers for absolute positions and optical shaft encoders for wheel speed). Four light level sensors are placed at each corner of the base, two task-specific 'cliff sensors' point down from its front face, and four capacitive sensors are arrayed along the inside of the body shell providing sensing of direct human contact. In the head, stereo microphones (in the base of the ears) and stereo cameras (in the eyes) are complemented by a sonar ranger in the nose and an additional four capacitive sensors over the top and back of the head (behind the ears).

Peripherals are reached on an I2C bus from the 'spinal processor' (ARM Cortex M0), which communicates via SPI with the 'brainstem processor' (ARM Cortex M0/M4 dual core), which in turn communicates via USB with the 'forebrain processor' (ARM Cortex A8). Division of the processing in this way is partly pedagogic and partly aesthetic, in service of the product's standard configuration, and plays no direct functional role. Nonetheless, it does align closely with the layered control architecture design (see Section 5.3.3). All peripherals and a level of control over processing are accessible from off-board through WiFi connectivity, and the forebrain processor is open if lower-level access is required (all processors can be re-programmed if desired, though with more onerous requirements to respect the specifics of the platform).

5.3.3 Control architecture and gross behaviour

MIRO's control system is a brain model with a layered architecture (Figure 5.3). That is, its most fundamental organising feature is the presence of sensorimotor loops layered on top of one another, so that lower loops function without the help of higher loops, but higher loops can modulate the behaviour of those lower down. Low-level loops implement reflex-like behaviours, immediate responses to sensory information that make use of neither memory nor signal analysis and can be implemented simply (for instance, soft threshold units respond to cliff sensor signals to inhibit forward wheel motion). Mid-level loops make use of short-term memory and within- and cross-modal signal relationships to implement 'hard-wired' behaviours that require co-ordination across motor systems (a major centre is a model of the superior colliculus that represents recent salient events in a multi-modal map of egocentric space and responds to specific 'innate' stimuli with directed action (Dean et al., 1989)). High-level loops use arbitrarily deep memory and inter-signal relationships to implement cognitive competences (reinforcement learning provides the ability to 'train' MIRO to perform simple stimulus-response tasks, for example).

Whilst this three-level break-down is simplified, it conveys well the architectural principle of layers of increasingly sophisticated processing, with each layer making an important contribution to overall behaviour rather than being obsoleted by higher processing. In order to arbitrate between behavioural sub-systems at mid and high levels a model of the basal ganglia (Gurney et al., 2001) is implemented in an abstract form as used in several of MIRO's predecessors (Pearson et al., 2007). Thus, MIRO's gross behaviour emerges from the competition between various sub-systems to explore locations with high sensory salience, escape from stimuli that are perceived as threatening, seek out goals (such as a charging station), have social exchanges with an interacting human, and so on.

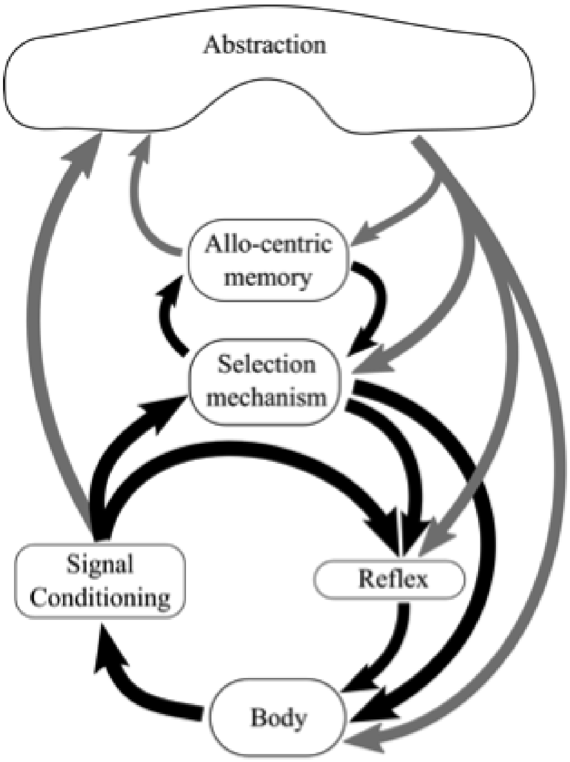


Figure 5.3. MIRO's brain-based layered control system. Loops at the lowest level are the least abstract, computing behaviours directly in the signal spaces of sensors and actuators. Abstraction, depth of memory, and complexity of computation all increase in progressively higher loops. Mid-level loops are focussed on action selection (including selection of spatial targets). Even the highest loops can modulate directly the behaviour of the lowest, as required (Collins et al., 2015).

5.3.4 Modelling and expressing affect

MIRO represents affective state using the circumflex model (Posner et al., 2005). This model represents emotions (as well as, on the longer term, moods and temperaments) as points in a space having dimensions of valence and arousal (Figure 5.4). These dimensions are purported to have neural correlates whilst terms used to describe emotions (such as 'excited') are cast as locations in this space. This stands in contrast to ‘basic emotions’ theory which considers individual emotions (such as ‘excitement’) to correspond to discrete neural systems. Whilst continuum models of this sort have overwhelmingly received attention in human studies, recently they have begun to be transposed into the domain of non-human animals (Mendl et al., 2010). These models are also remarkable for their clarity and accessibility for the non-psychologist, as well as for their light computational weight, and have, accordingly, received some attention from roboticists (Beck et al., 2010; Breazeal and Scassellati, 1999; Yilmazyildiz et al., 2013).

MIRO displays affective state through its behaviour. Affect is fundamental to MIRO’s functional behaviour because gross behaviours (such as approach, or flight) have unambiguous emotional correspondences and are, correspondingly, facilitated or suppressed by affective state. Affect is also communicated directly and explicitly through its encoding in MIRO’s non-locomotory movements. MIRO has mobile ears, eyelids, and tail expressly for the communication of affect, but body configuration movements are also driven by emotions (activation tending to lead to raised posture, for instance). Body language has been shown to be effective for the communication of emotions between humans (Wallbott, 1998) and consistent interpretation of the body language of animals by humans has been demonstrated (Wemelsfelder et al., 2012), though there is considerable variation between species in expression (Darwin, 2002). Moreover, the use of human-like body language in humanoid robots is effective for communication of emotion to naïve humans (Beck et al., 2010).

In addition, MIRO is equipped with six RGB LEDs (three on each side) under its body shell that can be controlled dynamically (at up to 50Hz; see Figure 5.5). Through these, MIRO can display arbitrary light patterns that change in parameters such as colour and rate in a bid to communicate affect. Whilst light displays offer rich expression and low cost, changing patterns of lights—in contrast to body language—do not have a direct biological analogue. Certainly, cultural associations exist for parameters such as colour—red/green for traffic lights is an almost universal contemporary code, for example—but reports have been presented of variability in these associations based on culture (Courtney, 1986), gender (Hurlbert and Ling, 2007), and context (Maier et al., 2009). There is a considerable literature reviewing the effect of colour on physiology, behaviour, and emotion, and individual and cultural differences in colour responses; some population relationships are present, but a clear picture has not emerged (Manav, 2007; Valdez and Mehrabian, 1994). Moreover, it is not clear in what way such associations would translate to perception of the affective state of a robot, nor whether these perceptions would be reliable in a naïve interactee. Work addressing this question to date has been somewhat informal and results variable (Haring et al., 2011). However, in order to ascertain if this non-mammalian channel of communication was behaviourally predictable by naïve users a study was conducted which is reported below.

5.4 Validating MIRO's Non-Mammalian Communication System

The following section presents a study conducted to validate one aspect of MIRO's design: the platform's ability to communicate affect through light. Ethical approval for this study was granted by the Ethics Committee in the Department of Psychology at The University Of Sheffield.

5.4.1 Method

In many cultures, red signals danger and green safety; red/white/green was therefore selected for encoding negative/neutral/positive valence. Red is also a signal for sexuality, and for the ripeness of fruit, and green for nausea and decay (the degree to which these associations are biological or cultural is not always clear), so this selection could equally well have been the opposite encoding; such observations underline the uncertainty in these associations and the need for empirical study. The rate of change of a light pattern may be intuitively linked to arousal—both breathing and heart rate, for example, increase in frequency with increasing physiological arousal—so slow/medium/fast was proposed to encode deactivation/neutral/activation (specifically, 0.25/0.5/2.5Hz, reflecting the frequency range of human breathing/heart rate). Thus, nine points in affect space could be encoded, in total.

The remaining parameters of a pulsating light pattern that could physically be presented through the three RGB LEDs available on each side of MIRO was arbitrarily selected. Specifically, the pattern at each parameter point was monochromatic, with sinusoidal intensity, and with a fixed phase offset between adjacent LEDs of $\pi/2$ radians. Whilst the pattern was chosen to be deliverable through MIRO's LED arrays, patterns were actually delivered to participants through a simulation of one of the arrays on a computer monitor. This choice reflects the more general nature of the study's question (is MIRO's light-based non-mammalian communication channel behaviourally predictable?), and was intended to eliminate possible sources of confound stemming from participants' perceptions of other aspects of MIRO's design and presentation (its shape, positioning, etc.). The actual colours delivered ranged, in each case, from zero intensity (black) to maximum intensity of either pure red, pure green, or white.

The methodology for measuring the effectiveness of these encodings for evoking emotional perceptions was similar to that established by (Beck et al., 2010). Naïve participants (n = 5, 2 female; M age = 30, SD = 5) were recruited informally from The University Of Sheffield Robotics Laboratory. Prior to study participation written informed consent was obtained from each participant. Participants were then asked to view simulated light patterns and indicate their perceptions of them on nominal and interval scales.

Participants were seated one at a time in front of a laptop computer. The experimenter gave them initial directions, and then left them to follow on-screen instructions. The computer displayed simulations of one of MIRO's light arrays (Figure 5.6) at the nine points in affect space comprising each possible combination of negative, neutral, and positive valence and arousal (for analysis, negative/neutral/positive were assigned the values -1/0/+1). Participants were first exposed, over the course of thirty seconds, to all nine points, with instructions to watch the patterns. They were then presented with each of the nine points again, in random order—these are referred to as the 'presented' affect values. Participants were asked to fill a response sheet for each presentation, comprising:

1. Which of the following words best describes your perception of the emotional state represented by the pattern of light? Please circle one:

   Happy – Depressed – Calm – Stressed – Relaxed – Sad – Alert – Upset – Elated – Nervous – Contented – Bored – Serene – Excited – Neutral – Tense

2. If you think another word or phrase better describes your perception of the emotional state represented by the pattern of lights please write it here: —

3. Place a vertical mark on the line to indicate your perception of the level of arousal represented by the pattern of lights, from relaxed to

aroused:

Relaxed ———————————————— Aroused

4. Place a vertical mark on the line to indicate your perception of the level of happiness represented by the pattern of lights, from unhappy to happy:

   Unhappy ———————————————— Happy

The terms used in question 1 were taken from Posner et al. (2005), with the addition of 'neutral', following Beck et al. (2010), and presented in a randomised order. At the end of the response phase the experimenter conducted a short informal interview in which participants were asked whether they found the question 1 word list adequate. If the participant had answered any question 2 with a word or phrase of their own this was also discussed. The interview was conducted to establish whether the participants had perceived the patterns in emotional terms at all and, if so, whether the word list had allowed them to express their perception. At the end of the interview participants were debriefed.

For numerical analysis, numerical values in $[-1, +1]$ were allocated for valence and arousal with each of the terms used in question 1 (each taking a position in affect space on the unit circle, as indicated by their location in Figure 5.4) and with each of the marks in questions 3 and 4 (with the left/right extrema on the scales being transposed to -1/+1). These values, recovered from participants' responses, are referred to as the 'reported' affect values. Analyses of the reliability of the relationships between presented and reported affect values were conducted independently for valence and arousal.

5.4.2 Results

Initially the results pooled across participants were analysed; these results are graphed in Figure 5.7. Positive correlations were identified between presented and reported values for both parameters when using both approaches to reporting. The relationship was apparently robust in all four cases, with between 25% and 70% of the variance in reporting explained by a simple linear predictive model.

The identified relationships were then exploratively reviewed on an individual basis (see also Figure 5.7). Data from each participant displayed relationships of the same polarity as those displayed by the pooled data, indicating that pooled results reflected the responses of all participants in this sense.

In response to question 2, only two responses were received (of a possible 45). These are the terms indicated in typewriter font in Figure 5.4, and they are placed in the affect space at the location of the presented stimulus for each of those 2 trials. Informal interviews generally indicated a high level of satisfaction with the word list for expressing participants' perceptions.

5.4.3 Discussion

MIRO has been developed to mimic biological communication channels, in particular those used by mammals. In addition to these, MIRO also uses an expressive modality via coloured lighting patterns. Dynamic lighting patterns may not have direct biological analogues (though see cephalopods (Mäthger and Hanlon, 2007)), but colours are strong situational signals (being indicative of the presence of ethologically-relevant items such as blood and food), and rate of change may be associated with physiological markers of arousal; colour also has cultural associations, which may be more or less reliable depending on participant and context.

As the focus of Chapter 7 is to be the importance of MIRO's predictable mammalian-like behaviour it was important to establish whether MIRO's additional non-mammalian communication channel was behaviourally predictable by naïve users, much as a wagging tail is indicative of an emotionally engaged and 'happy' agent. The results of the work presented in this chapter support the hypothesis that this non-mammalian channel is predictable by demonstrating that patterns of pulsating lights can evoke reliable perceptions of affect in naïve participants.

Results from individual participants were suggestive of consistency, at least at the grossest level, in the selected participant group (participants were selected opportunistically in a British laboratory, and cultural background was neither recorded nor used in participant selection). However, it should also be noted that the results are suggestive of some differences between the four analyses (nominal/interval, valence/arousal) in the variability both between individuals and between reported and presented affective states—in particular the proposed encoding for valence seems to be more effective than that for arousal. For the purposes of the work presented in Chapter 7 this information is not directly relevant, however it is something that could be addressed by future research and development work on the MIRO platform.

In sum, the MIRO platform, though furless unlike PARO, has an advanced affective control system based on a scientifically established understanding of mammalian, and non-mammalian, communication and behavioural patterns. In addition to its behavioural channels MIRO offers its user a colour communication channel which provides a reward system for its user when haptically manipulated. Though this feature is not directly relevant to the subsequent hypotheses of this thesis it should be noted that its existence has been taken into consideration throughout the experimental phases.

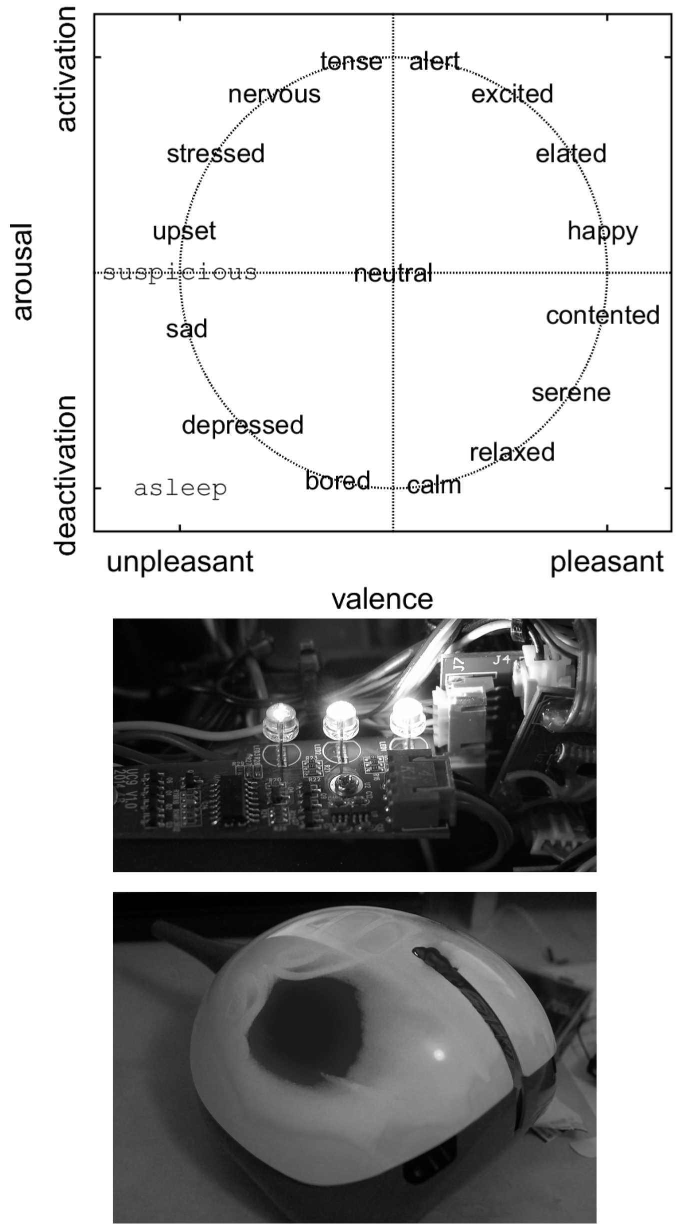


Figure 5.4. (Left) Circumplex model of affective state is a space with valence and arousal dimensions. Names for states (sans serif font) are taken from Posner et al. (2005), except for two suggested by experiment participants (typewriter font, described below). (Top right) One way in which MIRO expresses affect is through a changing pattern of coloured lights. (Bottom right) False colour image of one of the lights as it appears through MIRO's body shell.

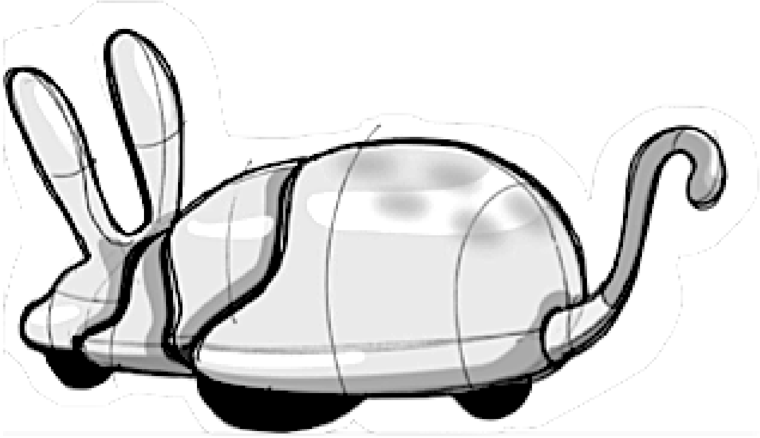


Figure 5.5. MIRO is equipped with six LEDs under its body shell through which it can display arbitrary light patterns in a bid to communicate affect: Prototype artwork showing original conception of signaling LEDs.

Figure 5.6. Stimulus presentation tool. Stimuli ($N = 9$) were presented in random order for each participant, who clicked NEXT when ready to move on.

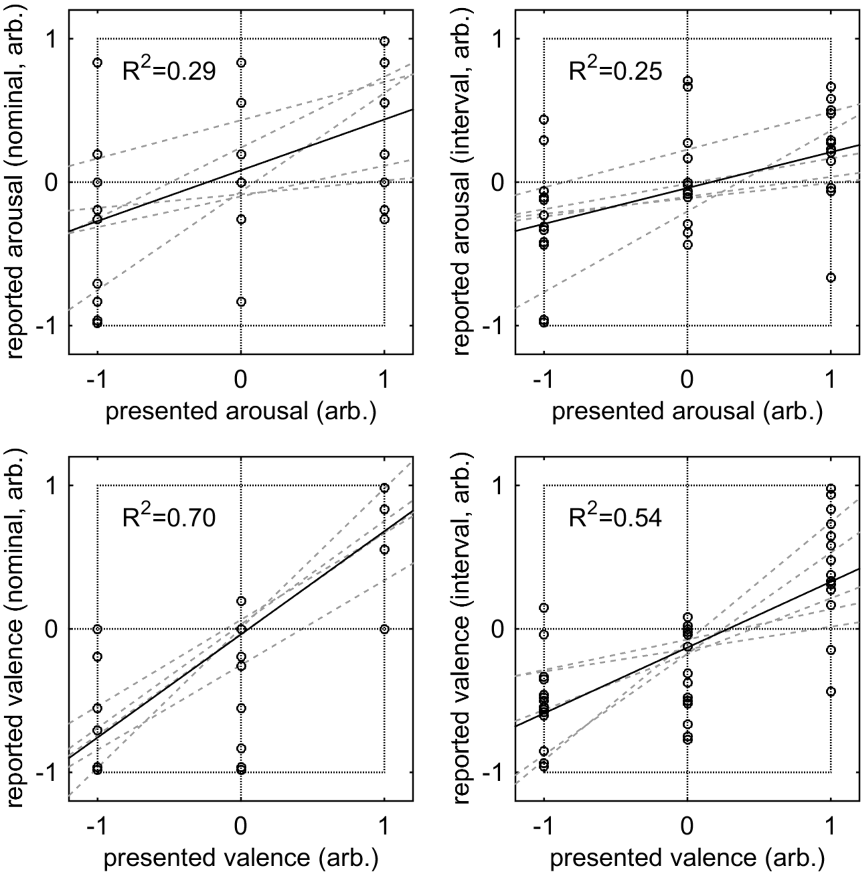


Figure 5.7. Reported affect values against presented affect values. Top/bottom: arousal/valence. Left/right: nominal/interval reporting. All units arbitrary (arb.). Individual trials (circles, $N = 45$). Trend line (solid) and R^2 values are from simple linear regression of pooled data ($N = 45$). Trend lines over samples from each participant ($N = 9$ per participant) are also shown (dashed grey).

Chapter 6

Cultural Effects and Translating the Felt Security Scale

6.1 Summary

In Chapter 4, engagement with PARO was established as being linked to both felt security and physical interaction with the robot, independent from the measures of individual differences taken in the same study. However, the large variation between individuals in how they respond to the same robot indicate that individual differences do play a part in how users engage a robot, and as such should be considered important as underlying the willingness of a user to engage with a social biomimetic robot. An individual's mental model of a robot, which drives the extent to which they wish to engage with it, as well as the level of expectation, acceptance and believability they will have of the machine, is in part constructed via the social culture in which an individual has been brought up and lives in. In this way it seems that culture, much like individual differences, will be important when considering acceptance, use, and the development of robots intended for RAT.

In Chapter 7 the role that culture might play in the acceptance and believability of MIRO will be explored, and measured via the impact that robots programmed with different mammalian-like behaviour have on a user's well-being. The introduction of Chapter 7 will be a discussion of how users might be expected to respond to platforms which exhibit predictable or incongruent behaviour. Before that discussion, and the presentation of the cross-cultural study, Chapter 6 will address the question of the effect of culture on robot engagement. Can RAT be expected to work equally across cultures, given what is known about different country's cultural attitudes towards robotics? Perhaps opinions regarding robots vary a great deal between individuals because individuals may possess limited mental models of robotic units, both as experienced via the media and in real life, or because culturally a more general pattern exists, reflective of how different cultures present robots to populations as a whole.

In order to explore this question of the potential effect of culture on RAT, Chapter 6 will explore established literature on different cultures and the different levels of exposure populations have to robots through the media or personal experience. These different levels of exposure are in turn thought to influence how sensitive individuals from specific cultures are to robotic systems. Japan is often cited as an example of a culture that embraces robots, leading to the assumption that Japanese citizens are more sensitive to robot interactions as a result. Thus Chapter 6 will focus predominantly on Japan. This is also reflective of the fact that the cross-cultural study to be reported in Chapter 7 was conducted in the UK and Japan.

In order to conduct a study in Japan that measured user well-being, the English FSS was translated into Japanese. As such, Chapter 6 also presents the translation of the Felt Security Scale into Japanese to produce the new JFSS instrument. This is presented with a discussion of the difficulties involved in translating psychological instruments, and the importance of semantic equivalence. A total of 41 participants completed the scale, as well as the Japanese version of the ECR which served as a measure of convergent validity. The JFSS showed excellent internal consistency (Cronbach's α = .871), and significantly negatively correlated with the Japanese ECR as expected (Pearson's $r(41)$ = -.36, p = .010 (anxiety), and $r(41)$ = -.32, p = .021 (avoidance)). Although the participant sample used were not representative of a clinical sample, such as that which would receive RAT, the translated instrument performed well. Thus, the production of this new instrument also demonstrates how transferable the comparative methodological approach of this thesis is to cross-cultural work.

6.2 Introduction

[P]eople's views of the world, of themselves, of their own capabilities, and of the tasks that they are asked to perform, or topics they are asked to learn, depend heavily on the conceptualisations that they bring to the task.

Donald A. Norman (Gentner and Stevens, 2014, p. 7)

It could be argued that culture should be taken into consideration when considering the implementation of RAT, and it should not be assumed that one method will work with equivalent success in all countries (as, for example, single methods of treating mental health are not always equally effective across different countries, see US Surgeon General (2001)). Although Chapter 4 did not significantly attribute differences in engagement with PARO to individual differences in caregiving or attachment style, it remains that there is large variation between individuals in how they respond to the same robot (Jones et al., 2008). Why these differences exist is not known. However, the existence of such differences is indicative that individual differences, of which cultural background is a part, play a role in how users engage a robot (see Chapter 4). Therefore, as with AAT, when the individual differences and cultural experiences of a client are taken into consideration when pairing them with an AAT co-therapist (e.g., as explained in MacNamara et al. (2015)), the individual differences and cultural experiences of users should also be considered important when considering the willingness of a user to engage with a social biomimetic robot. Given this, further exploration of why there are differences between individuals in how they engage a robot intended for RAT is required if researchers want to better understand the acceptance, use, and development of such RAT robots.

Differences between individuals' interactions with the world are related to differences between individuals' mental models of the world and its components (Denzau and North, 1994; Markus and Kitayama, 1991). As discussed in Chapter 3, this includes differences in interpersonal experiences across the lifespan which are internalised as cognitive models that guide behaviour and affect (Mikulincer, 1986). An individual's mental model of a robot drives the extent to which they wish to engage with it, as well as the level of expectation, acceptance and believability they will have of the machine. A mental model provides an individual with the means of predicting and explaining their environment and the artefacts therein. They allow an individual to understand interactions via naturally evolving models which are not necessarily accurate, despite providing functionality. Mental models are constrained by such as an individual's background, their previous experiences with the system in which they are interacting, and the limitations of the processing power of the mind itself (Gentner and Stevens, 2014). Aside from individual cognition and direct experience however, cultural models and social norms also have a powerful influence over structuring individuals' mental models of their environment (Mantovani, 1996). Whilst it must be stressed that individuals are not cultural clones (Beach, 1990), social context is nonetheless integral to the construction of the context within which an individuals' working understandings of their world – their mental models – are formed.

Given this, consider robots: it could be argued that currently within the UK accurate depictions of state-of-the-art (SoA) robotics, as well as robot platforms themselves, are not ubiquitous. In the 2012 EU-wide 'Public Attitudes Towards Robots' report, 60% of responders wanted robots banned from the care of children, the disabled or the elderly (The European Commission, 2012). Whilst according to the 2016 Robotics Business Review, the UK is currently home to only three of the top 50 robotics companies in the world (Carroll, 2016).

Conversely, Japan is a world leader in robotics, with robots contributing to many areas of society, including entertainment, healthcare and manufacturing (MacDorman et al., 2009). Globally, Japan has the second largest ratio of robots to manufacturing industry workers, with 323 robots per 10,000 workers in its factories (Prodhan, 2015).[1] Japan also dominates the manufacturing of robots themselves, holding 60 percent of the global market (Prodhan, 2015). Whilst on May 16th, 2015, Japanese Prime Minister Abe opened the 'Robot Revolution Initiative Council' with a call to the nation's corporate sector to, "spread the use of robotics from large-scale factories to every corner of our economy and society" (Fensom, 2015). Historically Japan has been a world leader in social robotics too, with development on the PARO robot starting as early as 1993 (AIST, 2006).

---

[1] Japan is only out-placed by South Korea, which has 437 robots per 10,000 manufacturing industry workers.

Whilst the media continues to hail Japan as a country with a ‘craze’ for robots (MacDorman et al., 2009).

Level of exposure to robots occurs not only through direct experience of them but via the media too. In Japan robots are frequently portrayed in fiction, and in games and on TV, but unlike in western cultural portrayals of robots the typical ‘robots will take over the world’ scenario is not common (Bartneck et al., 2007). Alternatively robots in Japanese media, such as manga, are not exclusively ‘evil’ but represent a variety of roles unrestrained by their technical nature (Bartneck et al., 2007). Kaplan (2004) argues that this phenomenon is due to the traditional differences between western and eastern culture’s analogies of machines. Western depictions are focussed on machines that challenge humanity’s ‘narcissistic shields’ by their very existence (Bartneck et al., 2007), whilst eastern depictions, in particular those in Japan, are framed by a culture which does not distinguish artificial entities from natural ones, but rather views machines as opportunities to understand the natural laws of humanity instead (Kaplan, 2004). With such complex nuances exposing populations to not only different degrees of tangible robotic technology but also different social norms outlining the concept of what a robot *is*, it may well be concluded that culture does influence individuals’ mental models of robots, and in turn how individuals take to robots on a personal level.

However, exactly how different cultural backgrounds might be impacting how individuals respond to robots is not as straightforward as simply stating that the Japanese love robots (Bartneck et al., 2007), whilst individuals from nominally western countries such as the UK or Australia are more wary (for example see Haring et al. (2014)). In a cross-cultural study of negative attitudes towards robots Bartneck et al. (2005) reported that in a sample of Dutch ($n = 24$), Chinese ($n = 19$) and Japanese ($n = 53$) participants the Japanese participants rated robots less positively than received opinion, that Japan is robot 'crazy', might have indicated. Instead the Japanese participants in the study expressed concern over the impact that robots might have on society. The study authors argued that high exposure to robot technology, aside from media portrayals, could be responsible for such an opinion. Exposure to real robots in recent years may have left the Japanese more aware of the actual abilities and constraints of SoA robotics. This begs the question of whether such a difference between individuals from countries that promote different levels of social exposure to SoA robots, would display different sympathies towards a robot that was behaving in a non-optimal manner.

Given the differences between the UK and Japan's socio-cultural exposure to SoA robotics then, perhaps individuals living in the UK would have different mental models of robots than those living in Japan. This in turn could lead Japanese individuals to have higher expectations of robot behaviour, than individuals from the UK. Assuming that individuals who are more used to robots would be less sympathetic to the robots' mistakes, than for those to whom robots present an entirely novel agent, might it be that individuals with more developed mental models of robots based on more experience with SoA robotics would treat MIRO differently? This idea will be explored in Chapter 7.

After a successful application to the Japanese Society for the Promotion of Science (JSPS) the opportunity to conduct this cross-cultural UK/Japan study arose. The open question of whether there really is a difference between Japanese and western European cultures' attitudes to robots, that extends all the way to the individual and their personal robotic preferences, was addressed by conducting an identical study in both the UK, with a western European participant group who all lived in the UK, and again in Japan, with a Japanese participant group who were all born, brought up, and currently live in Japan. In order to conduct this cross-cultural study the English Felt Security Scale needed to be translated into Japanese.[2]

6.3 Translating the Felt Security Scale (FSS)

6.3.1 Semantic equivalence

It is well established that a single forward and back translation procedure is an insufficient method of making and checking the quality of a translation (Van Widenfelt et al., 2005). The process of translation must take into

---

[2] This author is a fluent Japanese speaker so this step was deemed possible as part of the thesis process.

consideration semantic equivalence. Here, *equivalence* is when a pair of items, one from the original scale and one from the translated scale, are deemed to be measuring the same construct in equal proportions such that there is an equal probability of getting the same response from both items despite them being in different languages (Hulin, 1987). In order to achieve this a hierarchical procedure of translation, back-translation, and then verification is required to establish construct (semantic) equivalence (Brislin, 1970).

Although attachment and related constructs, such as felt security, have been criticised for being ethnocentric (Rothbaum et al., 2000), the ECR has been successfully translated into multiple languages including Chinese (Mallinckrodt and Wang, 2004) and Japanese (Nakao and Kato, 2004). This is because ideal translations take into consideration the fundamental need for semantic equivalence. A hierarchical translation process should incorporate not only an understanding of bilingualism from the translators, but also biculturalism. Such that the translators possess not only a knowledge of the instrument being translated, but are also very familiar with the target culture (Mallinckrodt and Wang, 2004). This is the most important process required when producing an adequate instrument translation in order to ensure semantic equivalence is truly established. As such translated instruments may ignore the idiosyncrasies of the original format, in order to incorporate necessary culturally specific adaptations. This is particularly the case when a culture is viewed as having a potentially significant impact on how concepts are expressed in various languages (Herdman et al., 1997). In the case of the presently translated FSS this meant taking into account the need for a translation that allowed for full sentences in order that the instrument's items were clearly contextualised. For example, eastern cultures have been reported to interpret 'self-esteem' (which the construct of felt security is in part comprised of) somewhat differently to western cultures (Wang and Ollendick, 2001), and understanding cultural nuances such as this was important when translating the FSS items in a way that made it clear to the Japanese individuals completing the instrument that the items it discussed where referring to them, as individuals, and not to others as might otherwise be unclear without the use of very specific sentences to convey each item. Thus, although limited by both time and resources in order to conduct as thorough a translation as would be desired (by, for example, validating the Japanese FSS with a participant sample numbering into the hundreds or thousands) the translation of the FSS conducted in Japan, and here reported, did take into consideration both the linguistic as well as the cultural differences required to convey the same concepts.

In translating the FSS into Japanese, having a Japanese version of the Adult Attachment Style Scale Experiences in Close Relationships Scale (Japanese ECR; Nakao and Kato (2004)) proved extremely useful as a way of beginning to understand how felt security might be translated, without losing semantic equivalence. Having the Japanese ECR facilitated the translation of the FSS by providing an idea of the terms that could be used to express similar concepts seen in both the English FSS and ECR.

For example, how to describe being cared for by another human as an adult (as opposed to an infant), and how to describe love (a term used with much nuance in Japanese).

Finally, it should be noted that the description of the translation process given here is done so in as much detail as possible without going into linguistic specifics. Although examples will be given to highlight the process, note that for the ease of the reader these will be written in *romaji*, the romanised Japanese alphabet, and not in Japanese *kanji* or *hiragana* scripts.

6.3.2 Producing the Japanese Felt Security Scale (JFSS)

The English FSS (Luke et al., 2012) is a psychological instrument for measuring felt security. This is a concept which comprises feelings of care, love and self-esteem, as well as safety (Bowlby, 1969; Holmes and Rempel, 1989; Mikulincer, 1986; Murray et al., 2000). The original instrument was reliably validated by Luke et al. (2012): Cronbach's $\alpha = .97$; $M = 4.25$, $SD = 1.39$.

In the first step of translating the FSS into Japanese a native Japanese speaker (Professor Yoichiro Yoshikawa of Osaka University) and a native English speaker (this author) reviewed all the terms of the English FSS: *comforted, supported, looked after, cared for, secure, safe, protected, un-*

threatened, better about myself, valued, more positive about myself, I really like myself, loved, cherished, treasured, and adored. The concepts were then discussed as they are understood within Japanese, both linguistically and conceptually. Both translators were English-Japanese bilinguals, and also bicultural individuals who had extensive experience of the target culture (one being a native Japanese, the other having lived in the country for five years and self-reported as very familiar with the Japanese culture); as well as having spent substantial time living outside their native countries (in America and Japan respectively). The native British translator (this author) also had a thorough knowledge of the English FSS, which allowed the translators to work beyond merely having knowledge of the two languages, by also having an in-depth knowledge of the subject matter of the instrument.

Direct translation of each item was not possible due to several reasons. For example, it is difficult to translate the item *protected* whilst retaining semantic equivalence with the English. Directly translated into Japanese, *protected* is *hogo*, a word used when referring to children or an individual in need of social support (such as someone who is very ill). It would be unusual for an adult, speaking about themselves, to rate themselves as requiring protection at all, given the Japanese cultural ideal that adults should not need to be protected. Further, words which conceptually group together in one language do not necessarily share the same conceptual family in the other. For example, in the English FSS, *cherished*, *treasured* and *adored* are all separate items, although they respectively cover the same conceptual idea: to protect or care or value someone lovingly. However, in Japanese these terms are difficult to translate as individual items. For example, the single idea of 'value' covers both being *cherished* and *treasured*, whilst *adored* is considered close to simply being loved by another human.

Given this issue the original 16 items were initially translated into an alternative English sentence that more clearly described the conceptual meaning that the item was conveying. In doing this some of the 16 items were grouped together, or placed in more than one group. Thus, for example, the original FSS items *treasured* and *cherished* were grouped together and translated as *I feel like an invaluable person*, whilst *cared for, valued* and *cherished* were also grouped together and translated as *I feel I am an important/valued person* (see Table 6.1). Note in this last translation that 'important' and 'valued' are given equal weighting in the translated sentence, this is because the Japanese equivalent of the concept, *taisetsu*, means both 'important' and 'valued'. Understanding the Japanese version of each of the secondarily translated English concepts was an important part in the translation process, and required bicultural understanding in order to be fully realised. One item, *better about myself* was not included in this stage as it is difficult to convey this concept in Japanese without a temporal explanation. However, there does exist a short version of the English FSS (Luke et al., 2012), comprising 10 items: *comforted, secure, supported, safe, loved, protected, better about themselves, encouraged, sheltered, and unthreatened.* Internal consistency was measured by Luke et al. (2012), who reported high reliability: Cronbach's $\alpha = .97$; $M = 4.17$, $SD =$

1.53. From this shorter 10-item FSS *encouraged* and *sheltered* were taken and included in this stage of the translation process. See Table 6.1, for the 15 alternative English sentences that were produced at this stage.

Table 6.1. The original 16-items of the FSS (excluding *better about myself*), plus two extra items from the 10-item FSS (*encouraged, sheltered*) were conceptually grouped (with some overlap of items) and modified to create 15 new English sentences which more clearly described the concept being conveyed by each item. This was an important step to take prior to translating the items into Japanese in order to maintain semantic equivalency.

Item from English FSS (Single or Grouped)	Secondary English Translation
Comforted	I feel comforted
Supported	I feel supported
Looked after	I feel looked after
Cared for/Valued/Cherished	I feel I am an important/valued person
Secure (reassured)	I feel I am reassured
Safe	I feel safe
Protected/Sheltered	I feel I am protected/sheltered
Loved	I feel loved
Treasured/Cherished	I feel like an invaluable person
Loved/Adored/Cherished/Esteemed	I feel like I am dear to someone
That I like myself	I like myself
Sheltered/Positive about myself	I feel comfortable depending on others
Encouraged	I feel I can ask for help without resistance
Unthreatened	I do not think my environment is threatening

The secondary English translations, listed in Table 6.1, were then further reduced via another grouping method, in which sentences that when translated into Japanese were almost identical, were reduced to a single item. For example *I feel I am reassured* and *I feel safe* are conceptually similar, therefore only a single version of the Japanese translation of each was a chosen as the new item. In this case that was *anshin* (safety), which was used to create a new single item.

Finally the remaining sentences were translated into full Japanese sentences for the JFSS. Unlike in the English FSS, where single words are used to describe the terms, in the JFSS it was necessary to use full sentences to convey the concepts within a context. This was to ensure that the final JFSS would be readily understandable by all users. It can be difficult when using adjectives in Japanese to understand their meaning outside of a specific context. Therefore the use of full sentences in the JFSS allowed the instrument to be read as clearly indicating that each item was referring to how the reader felt at the time of completing the instrument. It should be noted that the final item-sentences were constructed in such as way as to ensure that no double-barrelled items were included, such that no single item contained two or more components that could warrant separate responses.

This resulted in 11 sentences which together conveyed the concept of felt security. In the English FSS each item was listed with its corresponding negative version, thus for the JFSS each item-sentence has a negative equivalent. See Table 6.2, for the English translations of the Japanese sentences, and Figure A.2 in Appendix A, for the Japanese version.

Table 6.2. The final 11-items of the Japanese Felt Security Scale (JFSS), full positive and negative sentences are given. Concept being conveyed by each item in italics. Notes in brackets indicate information explicitly conveyed by the Japanese translation of the sentence.

Positive JFSS Item-Sentence	Negative JFSS item-Sentence
I feel that I am being *supported*	I do not feel that I am being *supported*
I feel I am *cared for/looked after*	I do not feel I am *cared for/looked after*
I feel *secure* (subjective sense)	I do not feel *secure* (subjective sense)
I do not feel that I am *threatened*	I feel that I am *threatened*
I feel that I am *loved*	I do not feel that I am *loved*
I feel I am an *important* person	I do not feel I am an *important* person
I *like myself*	I do not *like myself*
I am *not hesitant to ask other people for help* (reassured)	I am *hesitant to ask other people for help* (not reassured)
I feel I am *valued/treasured/cherished*	I do not feel I am *valued/treasured/cherished*
I think I am *respected/adored/esteemed*	I do not think I am *respected/adored/esteemed*
I think I am *encouraged/supported* (to do things)	I do not think I am *encouraged/supported* (to do things)

The 11 items of JFSS sentences (22, including the negatives) were then passed to three Japanese bilinguals, who though native to the target culture had spent substantial time in English speaking countries. They were also naïve to both the FSS and its intended use. These individuals back-translated the Japanese sentences into English sentences they felt best expressed the concepts being described in Japanese. There was almost no variation between the three back-translations received from the secondary translators. Through this process it was established that the Japanese translations were good matches for the original English.

Finally, in order to fully establish face validity, one more individual, who was not associated with Osaka University (where the scale translation was taking place), and who was neither an English speaker, nor someone who had spent any substantial amount of time outside of Japan, was given the JFSS and asked if they could completely understand the form, including the Japanese translation of the instruction[3]. This volunteer had no problem at all understanding the JFSS, and was able to complete it in a manner they felt truly reflected how they felt at the time. At this stage as face validity had been fully established it was concluded that the translation was effective, and the JFSS was ready to be used in the experiment being presented in Chapter 7.[4] Results from the administered JFSS were subsequently used to assess the instrument's internal consistency, as well as its performance adequacy via a test of convergent validity against results from the Japanese Adult Attachment Style Scale Experiences in Close Relationships Scale (Japanese ECR; Nakao and Kato (2004)). In confirming

---

[3] In English the Japanese instruction translates as: *Place a vertical mark on each line in a place that best indicates how you feel right now.*

[4] Due to restrictions on time available in Japan to conduct experimental work and access to limited resources, the JFSS had to be used in the experimental phase prior to confirming the instrument's successful translation. However, having had extensive experience using the English FSS, this author, as the experimenter in Japan, was able to judge the performance of the JFSS as it was being used. As results comparable to those obtained previously when using the English FSS in the UK began to emerge it became apparent that the instrument was working as expected, and confidence was high that the subsequent validity results would demonstrate that.

these the instrument's use in Chapter 7 was deemed appropriate.

6.3.3 Results

41 participants (20 female; *M* age = 19.9, *SD* = 1.7) recruited from Osaka University via the Intelligent Robotics Laboratory volunteer email list completed the JFSS. All participants were healthy, with no known physical, auditory or visual impairment. Prior to study participation written informed consent was obtained from each participant.

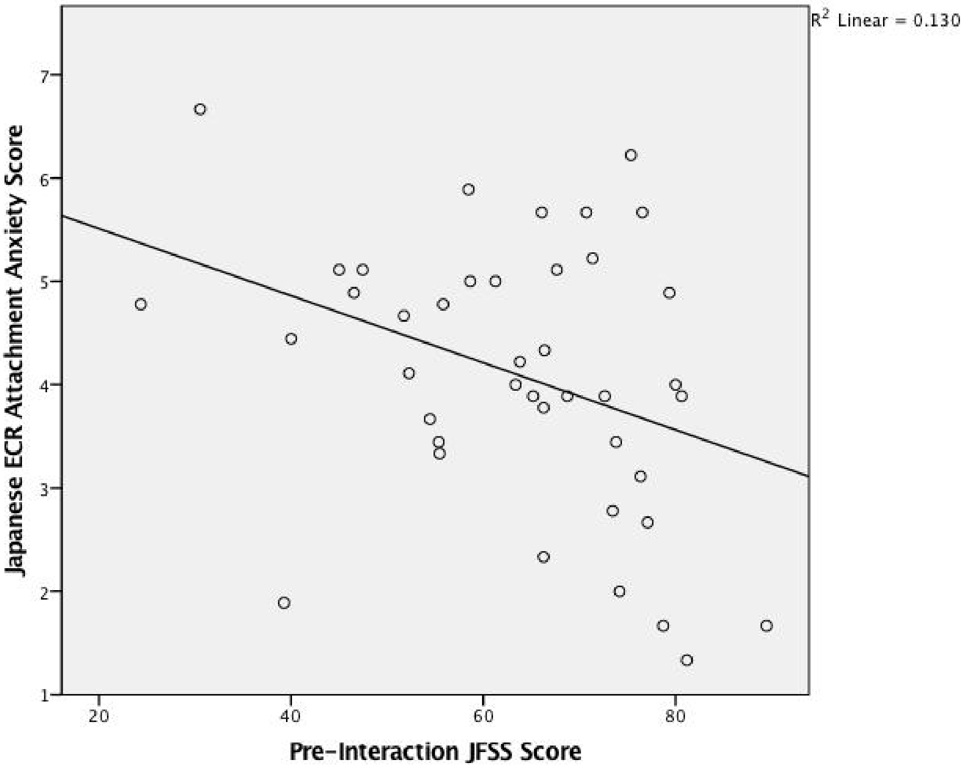


Figure 6.1. Scatterplot showing negative correlation between pre-interaction JFSS scores and attachment anxiety Japanese ECR scores.

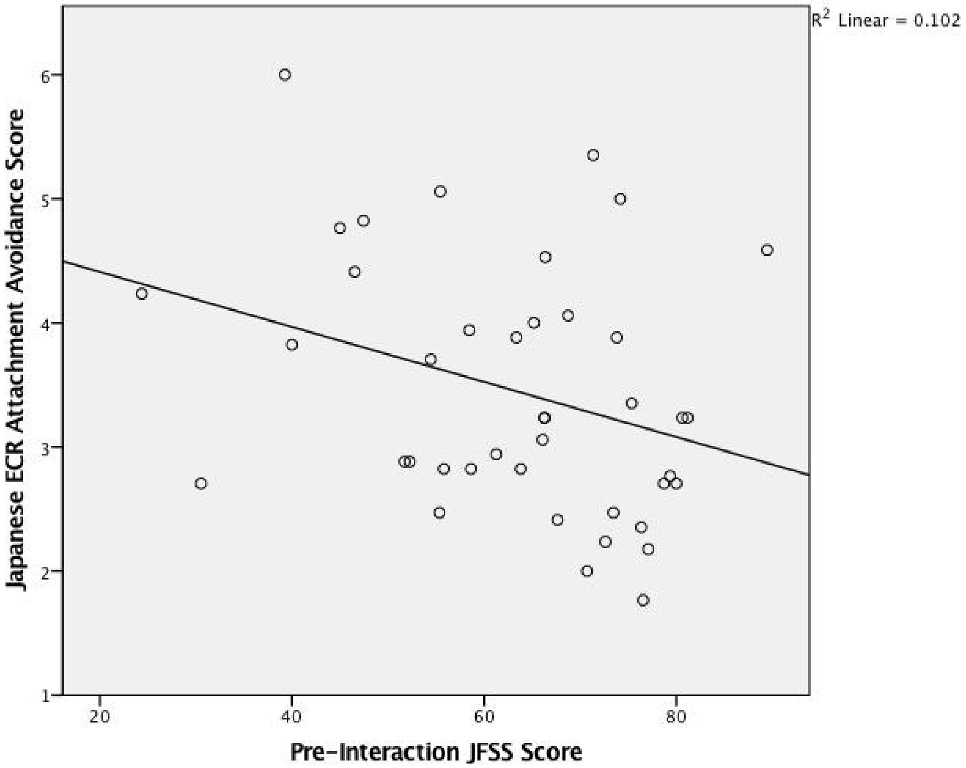


Figure 6.2. Scatterplot showing negative correlation between pre-interaction JFSS scores and attachment avoidance Japanese ECR scores.

Total scores for the pre-interaction JFSS items were computed, and internal consistency was confirmed, with high reliability, Cronbach's $\alpha =$ .871.

Participants also completed the 26-item Japanese ECR (Nakao and Kato, 2004). Results from the Japanese ECR produce two scores, one of attachment anxiety and one of attachment avoidance. As both of these dimensions are tapping into two forms of insecurity these sets of scores can be used to assess the performance adequacy of the JFSS via convergent validity, as both scales should negatively correlate with each other. After computing a Pearson product-moment correlation coefficient to assess the relationship between JFSS scores and the attachment anxiety and attachment avoidance Japanese ECR scores, negative correlations between both sets of variables was found, Pearson's $r(41)$ = -.36, p = .010 (anxiety), and $r(41)$ = -.32, p = .021 (avoidance). Scatterplots were produced to summarise the results (Figures 6.1 (anxiety) and 6.2 (avoidance)). Overall, there was a strong negative correlation between JFSS pre-interaction scores and Japanese ECR scores, such that the larger a participant's pre-interaction JFSS score, the smaller the participant's Japanese ECR scores.

6.3.4 Discussion

Although the participant sample used to assess the JFSS were not representative of a clinical sample, such as that which would receive RAT, the translated instrument performed well. Therefore, this successful translation of the FSS into Japanese has provided a new instrument for measuring the performance of robots intended for RAT. The production of this new instrument also demonstrates how transferable the comparative methodological approach of this thesis is to cross-cultural work. Using the JFSS, as well as the Japanese ECR, it would now be possible to repeat the work reported in Chapter 4, with PARO, in Japan.

The cross-cultural study reported in Chapter 7 uses the selectively modifiable robot, MIRO. Chapter 7 will explore the extent to which predictable mammalian-like behaviour is required from a robot to produce favourable outcomes in felt security, such as would be useful in a therapy scenario for improving client well-being. It is a two-part study, conducted in both the UK and Japan.

Chapter 7

MIRO in the UK and Japan: Felt Security and Behaviour

7.1 Summary

A companion robot that inspires in its user a sense of well-being would be ideal for implementing the goals of RAT: providing users of the robot with the sorts of positive benefits seen with animal ownership (as discussed in Chapter 2). In Chapter 4, it was established that a robot intended for RAT (PARO) can lead to positive increases in a user's felt security, even after controlling for the attachment and caregiving style of the user. This result parallels literature which supports the notion that AAT is deemed effective across individuals with differing attachment styles. In Chapter 5 the robot platform MIRO was introduced, as was the importance of a robot which allows for selective modification in order to conduct controlled studies that focus on single variables at a time. Chapter 6 explored the importance of culture for understanding why there exist differences between how individuals engage with robots. It also presented the newly created JFSS. In Chapter 7 these ideas are combined, and an experiment is presented exemplifying the sort of tightly controlled, single variable focussed experiment being promoted by this thesis. To that end Chapter 7 explores mechanisms of effect believed to be important when understanding and implementing effective AAT that may be relevant to robots intended for RAT.

Whereas Chapter 4 focussed on what features of users might be important for facilitating positive increases in user well-being (by focussing on individual differences and physical engagement), Chapter 7 shifts its attention to features of the robot itself. Chapter 7 asks: Is it the predictable mammalian behaviour of a social biomimetic robot that leads to a positive impact on a user's felt security? Moreover, Chapter 7 places this question into a cultural context based on extensive literature which addresses Japan's relationship with robotics and in which robots are described as being more abundant in everyday social culture than in western European countries like the UK.

The specific content of the chapter is as follows: continuing from a discussion in Chapter 5 asking what features of a robot might be important for promoting emotional engagement with a user, Chapter 7 begins with a discussion of the importance of behavioural predictability for encouraging positive engagement with another agent. Perception is then linked to the concept of mental models of robots as discussed in Chapter 6 to introduce the cultural aspect of the work presented in this chapter. Chapter 7 hypothesises that the more predictable the behaviour of a robot is, the more a user will engage with it, and this will result in greater levels of felt security as measured post-interaction. Further, Chapter 7 also hypothesises that cultural differences (as discussed in Chapter 6) may result in an observable difference between the felt security scores of the two country's participant groups (UK and Japan).

Results from across all conditions showed a significant positive change in felt security delta. However, contrary to the hypothesis concerning predictability, there was no effect of robot condition on this change in felt security. There was an effect of country on the baseline felt security scores, with Japanese pre-interaction felt security scores being significantly lower than UK pre-interaction felt security scores. However, despite existing literature claiming that the Japanese respond to robots in a markedly different manner to the British, no difference was found in how Japanese or British participants responded to a MIRO robot in either robot condition.

Chapter 7 begins to offer answers as to what it is about a social biomimetic robot that impacts the well-being of a user, as measured via felt security. In this case, across two countries with reportedly very different socio-cultural relationships with robotics there was no effect of predictable mammalian behaviour. However, this does not mean that behavioural consistency is not important, as the lack of a significant effect may be due to other factors. This result leaves many open questions about the mechanisms of effect driving positive psychological changes in users of robots designed for RAT. However, the methodology established in the work reported in Chapter 7 provides a promising route to addressing these. Finally, given the lack of a significant difference in felt security delta between the UK and Japan, Chapter 7 also raises interesting questions about the nature of received opinions regarding cultural difference, suggesting that the UK market for social biomimetic robots may be no less ready than the Japanese.

7.2 Introduction

Previous work exploring the impact on user satisfaction of an interaction with a robot designed to exhibit canid mechanisms of behaviour reported that neither the appearance nor behaviour of that particular robot is important for users when rating the robot's performance (Jones et al., 2008). The same study suggested that differences in individual preferences were so broad that they indicated the importance that individual differences play in overall robot performance assessment above and beyond the design of the robot itself.

The work described in Chapter 4, was in part an attempt to explore the variability in user reaction to a robot via a study of individual differences. Chapter 4 looked at a very specific set of characteristics important in social psychological understandings of human-human interactions: attachment and caregiving styles. It supported previous research, such as that by Jones et al. (2008), by once again demonstrating the existence of the wide variety of behaviours exhibited by a user interacting with a particular robot. However, Chapter 4 failed to attribute these differences to individual differences in attachment or caregiving styles, leaving the question of what design aspects in particular promote emotional engagement (EE) with a social biomimetic robot, such that a positive psychological change is experienced as a result, unanswered.

Continuing on from this, the mechanism of effect to be explored in Chapter

7 is the behaviour of the robot itself as. A lesson learnt from Chapter 4 was that there is a great deal of complexity involved in understanding the nuances of individual differences as being the root cause for varied and unpredictable responses to the same robot by different users. In order to find one mechanism, which could be directly understood as contributing to differences in the felt security delta of a user, a more precise study was required. It was decided that the present study should compare a specific feature of a robot that would allow conclusions to be drawn about the essential need for specific biomimetic features in RAT robots that are known to be important to the AAT process. Chapter 7 therefore explores the necessity for predictable mammalian behaviour in a robot designed for RAT, and how universal such a feature might be when considering cultural variability.

Humans can feel threatened and become stressed upon exposure to unpredictable behaviour (Jenkins et al., 1997), and behavioural stability is an important feature of an AAT co-therapist (certified AAT animals ensure that the animal can behave in a predictable manner before it is introduced to a care facility (Connor and Miller, 2000)). However, does it follow that because AAT animals are behaviourally predictable, robots intended for RAT must be predictable too, and in such a way as to be measurable via an observable difference in user well-being?

There is certainly evidence that behavioural predictably in robots is important in ASD therapeutic situations (Thill et al., 2012). This is due to the fact that robots can be programmed to communicate only relevant information to a child with ASD, avoiding the need to communicate unnecessary social information. For this same reason children with ASD are often very responsive to AAT animals who can be classed as “action agents” in contrast to humans who are “attitudinal agents” (Leslie, 1994): that is, animals communicate their intentions on a behavioural level that children with ASD understand more readily than intentions communicated by humans in social situations which contain large amounts of information whose relevance is difficult to decipher (Prothmann et al., 2009). However, beyond specific clinical examples the desire for behaviourally predictable robots, due to predictability being more effective at engaging a user, is widely reported (e.g., Breazeal and Scassellati (1999); Mutlu et al. (2009)), whilst robots that produce unpredictable or incongruent behaviours are seen as disturbing (Jones et al., 2008).

As explained in Chapter 5, in order to test the idea that behavioural predictability is important for increased levels of engagement with a biomimetically mammalian-like robot it is important to have an open platform that will lend itself to behaviourally controlled experiments. Therefore, given the flexibility of the system, its biomimetic morphology, and its relative size as comparable to a small dog, MIRO was chosen for this study.

7.2.1 Perception and performance assessment of robots

The public perception of robots is varied (e.g., Bartneck et al. (2007)). Robotic interfaces – in being both mechanical and novel – influence how the robot is perceived and received by a potential user. That perception will in turn influence how useful that unit is to its user for its specific role. In the case of a robot for RAT that role is to calm a user and to provide them with a consistent device through which therapy can be initiated in a manner similar to AAT (see Chapter 2). Given that, an ideal unit is presumed to be one which both produces expected behaviours and is also morphologically familiar without being too similar to its biological analogue as to elevate user expectations. Some of the literature surrounding this ideal has focussed on the limitations of biomimetic robots to produce a positive experience in a user. For example, humans can have adverse reactions to humanoid units for a variety of reasons. The height of a humanoid robot is important, with units that are too short (i.e. 0.6m) or too tall (i.e. 1.8m) heightening a user's anxiety (Hiroi and Ito, 2008). For humanoid robots its morphological familiarity also causes problems. Humanoid units are popular in science-fiction, and a user's knowledge of such advanced units results in a perceptual failure towards real robotic units, which do not meet fictional expectation (MacDorman and Ishiguro, 2006). For example, comparison between Star Wars' C-3PO, or any number of other realistic androids from science-fiction, can lead to confusion or frustration from a user towards the unsophisticated reality of state-of-the-art (SoA) humanoid units.

The acceptability of a biomimetic unit is in part reliant upon its visual and audial realism (MacDorman and Ishiguro, 2006), based upon the user's knowledge of the unit's biological counterpart upon which their mental model of the robot is then based. This mental model in turn facilitates familiar and predictable engagement with the robot. Robots that are biomimetic in their morphology, in the way they move, and that have expressive faces are immediately and intuitively engaging, owing to our familiarity with mammalian channels for conveying emotion and intent (Breazeal and Scassellati, 1999). Naïve 'users', for example, choose to interact to a greater degree with robots that include naturalistic body language in their interactions (Bruce et al., 2002), and robots can emit powerful social signals simply by following rules long-established by animals (Mutlu et al., 2009). On the other hand, knowledge that a robot is not biological does not eliminate the impact of these design strategies (Banks et al., 2008); and some evidence exists that even an explicit statement from a biomimetic robot's 'handler' that there is 'nobody home' leaves engagement more-or-less intact (Pearson et al., 2011). However, as with humanoids, knowledge of the biological counterpart can hinder the believability and acceptance of a zoomorphic robot. As explained in Chapter 4 (Section 4.3.1) PARO's seal-like morphology was chosen because of the expected lack of *a prior* knowledge of seals in its users, which was believed to aid PARO's acceptability. Theoretically it is a user's lack of preconceived ideas of the capabilities of a seal that leads to PARO being treated sympathetically. Whilst a general familiarity with mammalian channels for conveying emotion and intent aid in PARO's acceptability. This may also explain why the addition of incongruent features to a zoomorphic robot leads to self-reported dislike of the unit. For example, the addition of an artificial synthetic voice to AIBO and iCat to gets its user's attention (instead of barking or whimpering) has been reported as disturbing by users (Jones et al., 2008).

Arguably it is incongruence between the expected and the received that is relevant to the extent to which a user will find a social robot believable and acceptable. Thus the assumption is that believability and acceptability are important for the ease with which a user will interact with a unit. That ease in turn drives the extent to which that interaction will lead to positive changes in its user, which is a response expected from an interaction with a robot built for RAT. However, how true is this assumption? To what extent does a robot need to be believable and acceptable for a user to interact with it easily, such that the result of the interaction is a positive psychological change in the user? Further, how is 'believable' to be defined, such that it follows that a unit is 'acceptable'?

Bearing in mind the discussion in Chapter 5, Section 5.2, on the lack of well-controlled experiments in HRI which focus on single features at a time, as well as what was learnt from Chapter 4 about the usefulness of objective measures of psychological well-being for assessing interaction periods with a robot, the experimental design of Chapter 7 incorporated the following. Firstly, it was only focussed on exploring a single feature of the robot. Given the question of how 'believability' is to be defined, such that it follows that a unit is ‘acceptable’, the focus of the present study was on behavioural predictability (the details of which will be discussed below, in Section 7.3.2, which will outline the condition parameters that describe the mammalian-like behaviour performed by MIRO); acceptance of which was judged using a measure of psychological well-being. Thus, the second key design feature of the present study was its choice of performance measure.

Once again felt security was utilised to assess the impact of a robot’s performance on a user. Results from this measure of psychological well-being can be meaningfully translated to specific use cases, such as clinical RAT areas, in a manner that is harder to achieve using self-report measures given by healthy populations about a robot whose actual function is otherwise unknown to them (for a detailed discussion of felt security see Chapter 4, Section 4.2). Felt security is also related to a sense of energy (Luke et al., 2012), which constitutes a part of the resource remit required of an individual for environmental exploration. Secure relationships, which instil in an individual a higher sense of felt security in turn increase an individual’s felt energy, facilitating their desire to explore the environment. Relating this to EE, it could therefore be assumed that greater EE expressed between two agents (such as a user and a robot intended for RAT) would in turn be reflected in the user’s level of felt security.

Finally, perception of robot performance, and willingness to emotionally engage with a unit, may also be related to the culture of the user. As discussed in Chapter 6, although individuals from different cultures are exposed to different concepts of robots, and in turn are reported to have have different attitudes towards robots, it is difficult to express specifics when discussing culture in general (Šabanović et al., 2014). The importance of robotic technology to Japan has been nurtured by its government (Robertson, 2007; Šabanović, 2010). However, it must be noted that the received opinion which develops out of the nurturing of a cultural persona can obscure the fact that, regardless of cultural stereotypes, individuals will nevertheless form their own personal opinion of a robot when exposed to it based on their own mental model of the machine and what they expect from it. As was discussed in Chapter 3, Section 3.2.1, which noted that regardless of general maxims which play a role in how individuals interact with robots (see The Media Equation, (Nass and Reeves, 1996)), it remains true that individual differences also play a significant role in determining how an individual interacts with another agent (e.g., Caspi and Shiner (2006)). It is for this reason that such large individual differences are observed between individuals interacting with the same robot. Individuals are less sympathetic to technology that fails to meet their expectation: reporting frustration when technology that is difficult to use impedes working speed (Lazar et al., 2006). Such findings have prompted research exploring the divide that forms when expectation of technology, and anticipation of a goal, leads to frustration and lack of use upon failure to achieve the desired goal (Bucy and Newhagen, 2004; Nass and Reeves, 1996). From this it can be hypothesised that the more experience an individual has with SoA robotics, its limitations and its capabilities, the more likely that individual might be to notice when the robot does not behave as well as they believe it could. Thus, regardless of the personal-level individual differences observed between all individuals interacting with a robot, the socio-cultural environment in which an individual is born, brought up, and lives in can be expected to impact the level of their expectation of the robot, and the sympathy they will display towards to robot.

Given this, the experimental work reported in Chapter 7 is guided by the hypotheses in the following section.

7.2.2 Hypotheses

Given the seeming importance of behavioural expectation to the acceptance of a robot by a user, the focus of the present study was the extent to which a robot, exhibiting what is here described as ‘predictable mammalian behaviour’, was important to having a positive interactive experience with it as measured via felt security.

Level of engagement was constrained by the predictable reception the robot gave its user during the interaction period, with MIRO in Condition 1 (termed ‘congruent’) displaying congruent mammalian-like behaviour, and MIRO in Condition 2 (termed ‘incongruent’) displaying incongruent mammalian-like behaviour.

It was hypothesised that in both the UK and Japan, spending an interaction period with MIRO in Condition 1, a robot displaying predictably mammalian-like behaviour, would lead to greater EE and in turn larger felt security delta scores, than spending an interaction period with MIRO in Condition 2, a robot displaying incongruent behaviour.[1]

Based on existing cross-cultural literature such as that outlined in Chapter 6 describing the impact of culture on the formation of individuals' mental models, it was expected that the Japanese participants would respond less sympathetically than the UK participants towards a MIRO unit in Condition 2, which behaves incongruently, as the Japanese participants were expected to have mental models of social robots based on more experience with SoA robotics, which would lead to them having greater behavioural expectation from robots with which they are engaging. Thus, it was also hypothesised that in Japan the difference between conditions' felt security delta scores would be more pronounced than that observed with participants in the UK.

7.3 Method

A 2 x 2 two-way mixed measures design was used to explore whether engagement with a robot leads to an increase in user felt security. The two

---

[1] Felt security delta was treated as two separate scores in the analysis for Chapter 7: pre-interaction FS and post-interaction FS. However, both the single delta score and the two scores, pre and post, measure the equivalent change in FS.

between-subjects factors were: robot condition (congruent or incongruent) and country (UK or Japan); the within-subjects factor was the change in FS from pre- to post-interaction.

Ethical approval for both the UK and Japanese halves of this study was granted by the Ethics Committee in the Department of Psychology at The University Of Sheffield. For the Japanese half of the study this approval was further verified by the host department at Osaka University's Graduate School of Engineering Science; the Department of Systems Innovation, Intelligent Robotics Laboratory.

7.3.1 Participants

UK participants (n = 53, 27 female; M age = 32.1, SD = 9.67) were an opportunity sample recruited from The University Of Sheffield via verbal contact, or via email utilising the 'University Volunteer List'. Japanese participants (n = 41, 20 female; M age = 19.9, SD = 1.7) were recruited from Osaka University via the Intelligent Robotics Laboratory volunteer email list. All participants were healthy, with no known physical, auditory or visual impairment. Prior to study participation written informed consent was obtained from each participant. Participants in the UK were compensated for their time with a one-off payment of 5GBP. Participants in Japan were compensated for their time with a one-off payment of 2000JPY, as well as their travel expenses to and from the department. There were no instances of withdrawal. Participants were randomly allocated to one of the two conditions upon arrival at the experimental session. In the UK there were 26 participants in Condition 1 (congruent; 13 female), and 27 participants in Condition 2 (incongruent; 14 female). In Japan there were 20 participants in Condition 1 (congruent; 10 female), and 21 participants in Condition 2 (incongruent; 10 female).

7.3.2 Conditions

There were two conditions to the experiment. Each condition featured the same MIRO unit. Conditions were configured via a switch system located under MIRO's body shell and only accessible by manually taking the unit apart to reset the switches.

Condition 1: Congruent. Here MIRO's attention protocols are set to attend to both aural and visual stimuli by turning its 'gaze' to 'look at' a succession of chosen spatial targets. Thus, the overall behaviour is a 'series of orients'. Whilst in Condition 1 MIRO maintains an internal state of +ve/-ve valence which returns to neutral with a time constant of 30 seconds, reacts to external stimuli, and is expressed through various media. The stimuli that are reacted to are: 1) a head touch of any sort which drives valence +ve; 2) a body stroke head-to-tail which drives valence +ve; 3) a body stroke tail-to-head which drives valence -ve and; 4) loud noises which drive valence -ve. Expressions of valence state are: 1) tail movement, where a wag expresses +ve and a droop expresses -ve valence; 2) RGB LEDs under body shell which glow green to express +ve, red to express -ve, and remain white when valence is neutral and; 3) sounds are modulated by +ve valence upon which MIRO vocalises more trills and rising tones, and by -ve valence upon which MIRO vocalises more nasal, falling tones.

Condition 2: Incongruent. Here the MIRO unit was set to the same parameters as Condition 1, but with the following modifications. In the attention protocols the orienting system runs as in Condition 1 but the gaze target is overridden for each individual orient with a randomly chosen spatial location. Thus, MIRO's behavioural dynamics are fairly similar to those in the congruent condition, but there is no relationship between external stimuli and the robot's behaviour. The same dynamics were programmed for valence, but with a random choice of +ve/-ve. Thus, valence responds to the same stimuli, but unpredictably.

Checking detectable difference between conditions. It was important to establish the existence of a detectable difference between conditions such that regardless of the experiment's outcome it would be possible to state that the conditions were comparable, and contrastive enough to be recognisably different. In Condition 2, MIRO displays the same behaviour as in Condition 1 but without the intentional element, thus an individual engaging with MIRO should be able to tell whether MIRO is attending to them or not, and in turn may be more or less engaged in an interaction with it. In order to ascertain whether individuals naïve to the MIRO system could tell the difference between the congruent and incongruent system specifications, four individuals recruited as an opportunity sample from The University Of Sheffield, were exposed to MIRO and asked to describe its behavioural patterns over a period of five minutes, which was reflective of the time MIRO would be exposed to participants in the experimental phase. MIRO was exposed to each individual three times: the first time in congruent mode, the second time in incongruent mode, and the third time in congruent mode again. All four individuals could tell the first two modes apart, and correctly identified that the third mode was a repetition of the first mode. Modes one and three were described as, “responsive”, by all participants. Mode two was described as contrastively unresponsive. One individual described the mode two MIRO as, “shy”, whilst another participant described mode two MIRO as, “not wanting to look at me.” It was concluded that this brief check established the existence of a satisfactory level of detectable differences between conditions for the purposes of Chapter 7.

7.3.3 Measures

Felt security in the UK was measured using the 16-item Felt Security Scale (FSS; Luke et al. (2012)). The FSS administered in this study is presented as a visual analogue scale (VAS) where each question is answered by making a mark along a 100mm line on either end of which is an item relating to the full dimensionality of felt security being measured by the FSS (see Chapter 4, Section 4.5.1). Total scores for the felt security items were computed by Luke et al. (2012), who concluded that the items form a reliable scale, Cronbach's $\alpha = .970$.

Felt security in Japan was measured using the newly created 11-item Japanese Felt Security Scale (JFSS; see Figure A.2, in Appendix A), translated from the 16-item Felt Security Scale (FSS; Luke et al. (2012); see Chapter 6). As with the English version this measure was taken pre- and post-interaction with MIRO. The JFSS is also presented as a visual analogue scale (VAS) where each question is answered by making a mark along a 100mm line on either end of which is a sentence describing an item and its opposite relating to the full dimensionality of felt security being measured by the JFSS. Participants indicated how *supported, looked after/cared for, valued/treasured/cherished, secure, unthreatened, loved, important, respected/adored/esteemed, reassured, encouraged/supported, and I like myself* they were currently feeling. Total scores for the pre-interaction JFSS items were computed. The scale had high reliability, Cronbach's $\alpha = .871$.

Attachment style, was also measured in the Japanese sample in order check for convergent validity between the Japanese ECR and the newly translated JFSS. This was done using the 26-item Japanese version of the Adult Attachment Style Scale Experiences in Close Relationships Scale (Japanese ECR; Nakao and Kato (2004)), which is developed from the 36-item English ECR (Brennan et al., 1998). Results from the Japanese

ECR produce two scores, one of attachment anxiety and one of attachment avoidance. As both of these dimensions are tapping into two forms of insecurity these sets of scores can be used to assess the performance of the JFSS, as they should negatively correlate with each other. Participants responded to this measure on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*). Evidence for the reliability and construct validity of the Japanese ECR is rigorously made by Nakao and Kato (2004).

7.3.4 Procedure

In the UK the interaction session took place in the Human-Robot Interaction Lab at Sheffield Robotics, The University Of Sheffield. In Japan, the session took place at the Intelligent Robotics Laboratory, in the Department of Systems Innovation at Osaka University's Graduate School of Engineering Science, and all communication was held in Japanese.

Upon arrival to either lab participants were seated outside the interaction room and given paperwork to attend to whilst the experimenter went into the lab and turned on the MIRO robot and the lab's recording equipment as each session was filmed.[2] Participants were not initially informed that the session would be recorded so as to capture the most natural behaviours possible, uncompromised by participants potentially feeling nervous about

---

[2] Video data was collected to allow for future exploration of any significant effects of interactions.

being watched.

In the UK the paperwork included, 1) an information sheet which explained the study, whilst not mentioning that the interaction would be recorded; 2) the first written consent form to sign and; 3) a pre-session FSS (Luke et al., 2012) form to complete. In Japan the paperwork was, 1) a consent form to sign; 2) administrative paperwork for processing their monetary compensation and travel expenses and; 3) a pre-session JFSS form to complete.

When the participant completed their paperwork they informed the experimenter who then took the participant's paperwork and asked them if they had any questions. The experimenter then gave each participant the same scripted instruction.[3] The participant was then let alone into the interaction room and the door was shut behind them. In both rooms there was a table and chair. On the table was a MIRO robot set to either congruent or incongruent specifications. After five minutes interaction time the experimenter knocked on the door and entered the room. MIRO was then turned off and the participant was asked to complete a second FSS/JFSS.

The experimenter then left the room and asked that the participant open the door to let them back in once the second FSS/JFSS form was com-

---

[3] In English or Japanese depending on location: *Please can you turn your phone off. Please enter the room and sit on the chair in front of the table. The robot is on the table, and is already on. Please interact with the robot as you wish. Touch it and hold it as you like, explore its features, you can walk around with the robot if you wish. I will come back after five minutes and knock the door and ask you to fill in another form. Do you have any questions? Okay, if you need anything I'll be right outside.*

plete. At that time in the UK the experimenter took a seat at the table and explained to the participant that they had been covertly video recorded, and turned off the HRI lab equipment. All participants were then given an opportunity to have their recording deleted if they wished to withdraw from the study at that time. Participants were informed that they would still receive monetary compensation if they so wished to withdraw. The experimenter then gave the participant a second information sheet, which explained that the study was recorded in order that behavioural observations could be captured which could potentially be used in later studies. A second written consent form was then given to the participant to sign confirming that participants had acknowledged that their data would be securely kept by the experimenter. The experimenter then debriefed the participant fully with these two forms. The experimenter then handed the participant their monetary compensation, and finally led the participant out of the department.

In Japan the procedure at this point was slightly different. Once the second JFSS was completed the participant let the experimenter back into the room. The experimenter then asked the participant to complete two more forms. The first was a series of questions about the robot itself. These were presented as a visual analogue scale (VAS). Participants were asked to *Place a vertical mark on each line in a place that best indicates how you feel right now.* The questions were:[4]

---

[4] Note that the English versions of the three questions are direct translations from the Japanese.

1. The robot always responded to you — The robot did not respond to you at all

2. The robot had a mind like a living animal — The robot did not have a mind like a living animal

3. The reaction of the robot looked like one of an animal — The reaction of the robot looked like one of a machine

These questions were included to gather information on how the participants were perceiving MIRO which was of interest to the collaborators for future work. Question 1 was to find out if they could tell whether MIRO was being attentive or not. Whilst questions 2 and 3 were to find out how participants were perceiving MIRO.[5] The second form, the Japanese ECR, was placed face down on the table and the participant was instructed to turn it over and complete it once the robot questions were finished. Once these forms were completed the participant left the room and the experiment was over.

---

[5] Due to extenuating circumstance these questions were only asked to 28 out of a possible 41 participants: 14 in Condition 1 and 14 in Condition 2.

7.4 Results

All statistical analyses were run in IBM SPSS Statistics 23 for MAC OS X (10.10.5). The figures were created with the same software.

7.4.1 Tests for normality

The pre-interaction FSS and JFSS scores both had high reliability, Cronbach's $\alpha = .931$ and $\alpha = .871$, respectively.

Tests for normality were conducted on results from Conditions 1 and 2 separately for the the UK and Japan. Tests indicated that the pre-interaction FS scores for Condition 1 (congruent; UK $n = 27$, Japan $n = 20$) and Condition 2 (incongruent; UK $n = 26$, Japan $n = 21$) were normally distributed: z-skewness of -0.19 and z-kurtosis of -0.65 (congruent UK); z-skewness of 0.73 and z-kurtosis of -0.85 (incongruent UK); z-skewness of -1.23 and z-kurtosis of 0.47 (congruent Japan); z-skewness of -1.75 and z-kurtosis of 0.37 (incongruent Japan). Results from the pre-interaction FS score Kolmogorov-Smirnov test for the UK were, $D(27) = 0.108$, $p = .200$ (congruent), and $D(26) = 0.155$, $p = .109$ (incongruent), and for Japan were, $D(20) = 0.104$, $p = .200$ (congruent), and $D(21) = 0.146$, $p = .200$ (incongruent), indicating normal distribution in all samples.

See Figures 7.1 and 7.2 for histograms of the UK pre-interaction FS scores showing normal distribution, and Figures 7.3 and 7.4 for the Japanese. No data were excluded from subsequent analysis from either condition.

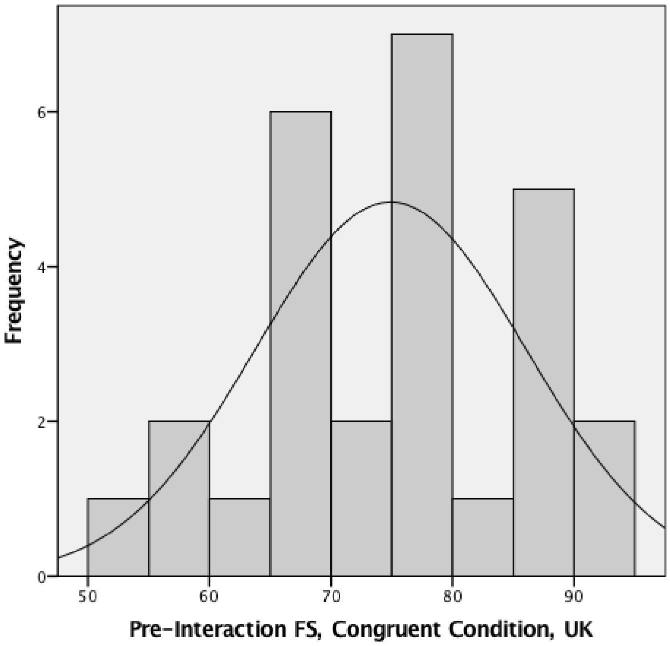


Figure 7.1. Histogram of the pre-interaction FS scores from the UK participants for Condition 1 (congruent) showing normal distribution.

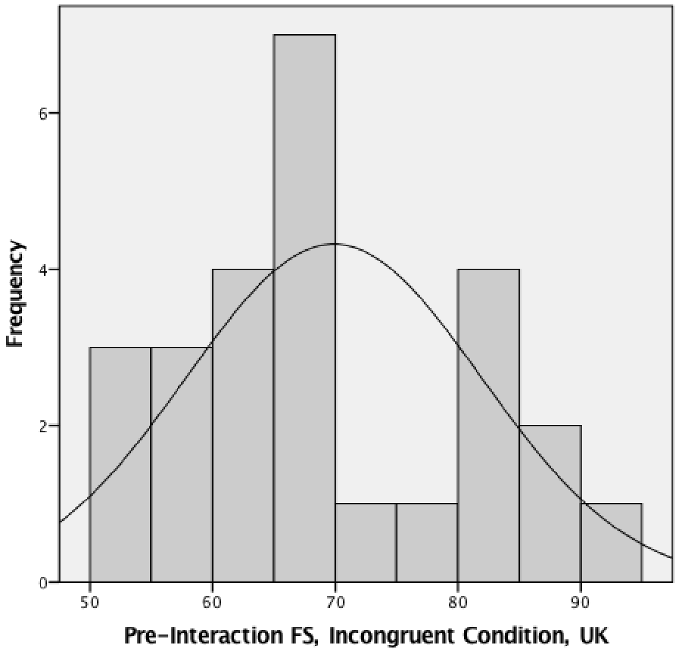


Figure 7.2. Histogram of the pre-interaction FS scores from the UK participants for Condition 2 (incongruent) showing normal distribution.

7.4.2 2 x 2 Two-way mixed ANOVA

A 2 x 2 two-way mixed ANOVA was conducted to explore both the main effects of condition, country and interacting with MIRO as measured by a difference between pre- and post-interaction FS scores, as well as interaction effects between those variables.

Results of Levene's test indicate the homogeneity of variance assumption was met for both pre- and post-interaction FS score varaibles: $F(3,\, 90) = 0.93$, $p = .429$ (pre-); $F(3,\, 90) = 0.18$, $p = .907$ (post-).

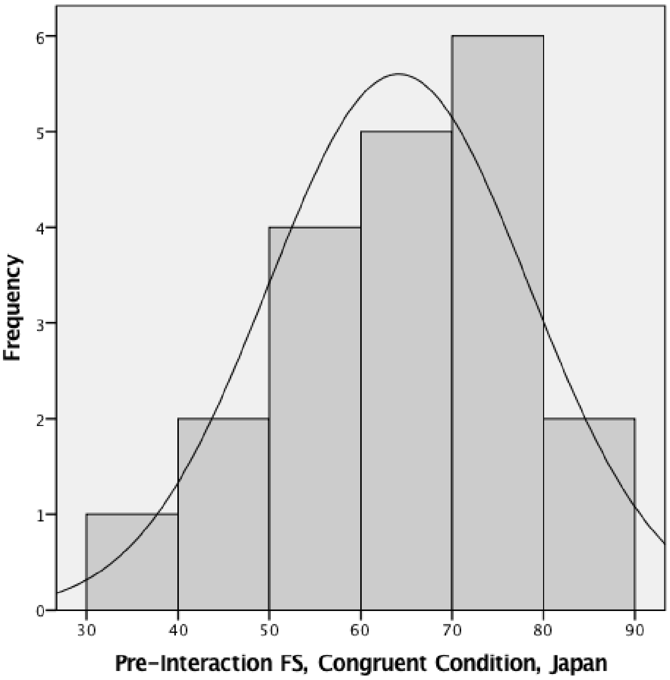


Figure 7.3. Histogram of the pre-interaction FS scores from the Japanese participants for Condition 1 (congruent) showing normal distribution.

Main effects

There was a significant main effect of FS score, $F(1,\ 90) = 8.51$, $p = .004$, $\eta_p^2 = .01$. This effect indicates that independent of the condition or country a participant was interacting with MIRO in, post-interaction FS scores were larger than pre-interaction FS scores (Figure 7.5).

There was a non-significant main effect of robot condition, $F(1,\ 90) = 1.16$, $p = .284$. Indicating that there was no significant difference in FS scores between participants interacting with MIRO in either the congruent or incongruent conditions (Figure 7.5).

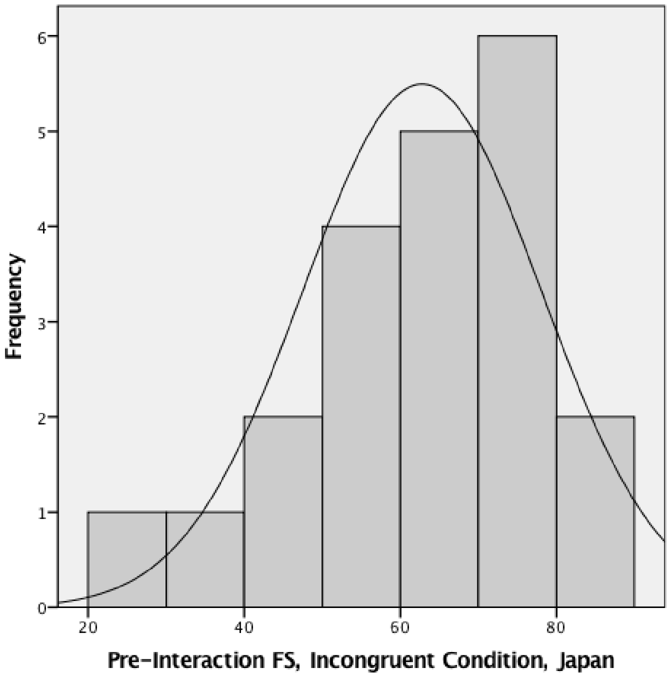


Figure 7.4. Histogram of the pre-interaction FS scores from the Japanese participants for Condition 2 (incongruent) showing normal distribution.

There was a significant main effect of country, $F(1, 90) = 12.13$, $p = .001$. This effect indicates that ignoring the condition a participant was in, pre-interaction FS scores were significantly different depending on whether the participant was in the UK or Japan (Figure 7.6).

Interaction effects

There were no significant interaction effects between FS scores and robot condition ($F(1, 90) = 0.09$, $p = .765$); FS scores and country ($F(1, 90) = 0.20$, $p = .654$); or FS scores x condition x country ($F(1, 90) = 0.03$, $p =$

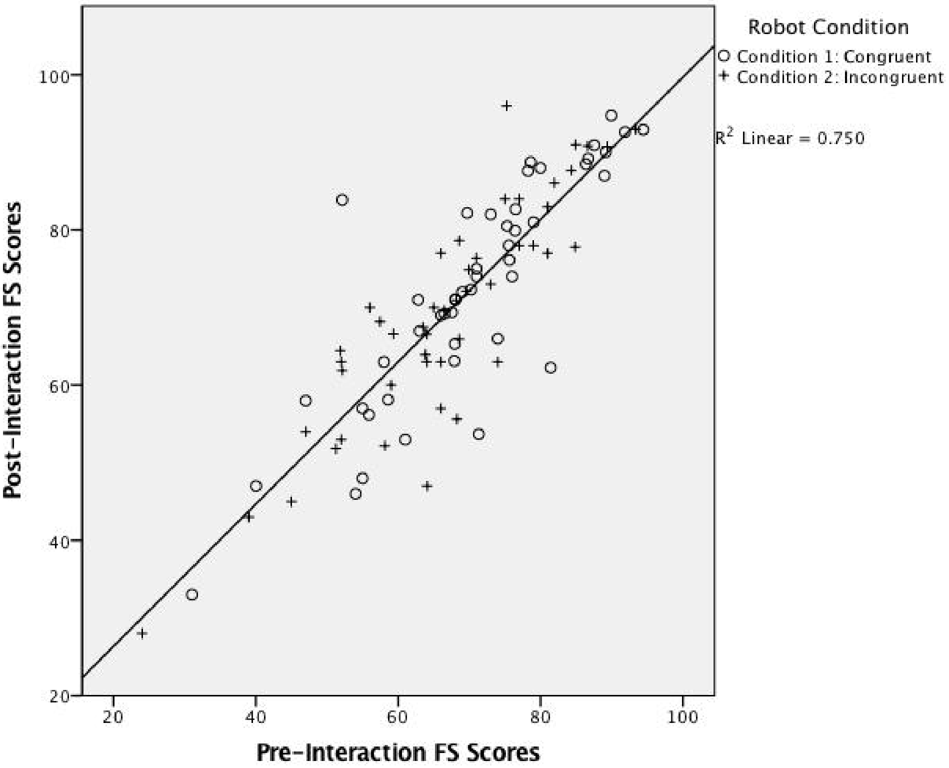


Figure 7.5. Scatterplot visualising pre- and post-interaction FS scores by robot condition (congruent/incongruent). There was a significant main effect of FS score, independent of robot condition.

.863).[6]

Taken together these results indicate that on average all participants saw an increase in their FS score post-interaction with MIRO and this was independent of the robot condition they were in or the country they were from. Although Japanese participants were significantly more likely to have lower pre-interaction FS scores than participants in the UK. See Fig-

---

[6] As the interaction results were non-significant no video analysis was conducted. Had there been a difference found in FS scores from pre- to post-interaction due to the condition or country manipulation, video analysis could have been conducted to look for differences in physical engagement across participants, which may have accounted for observed differences in the data.

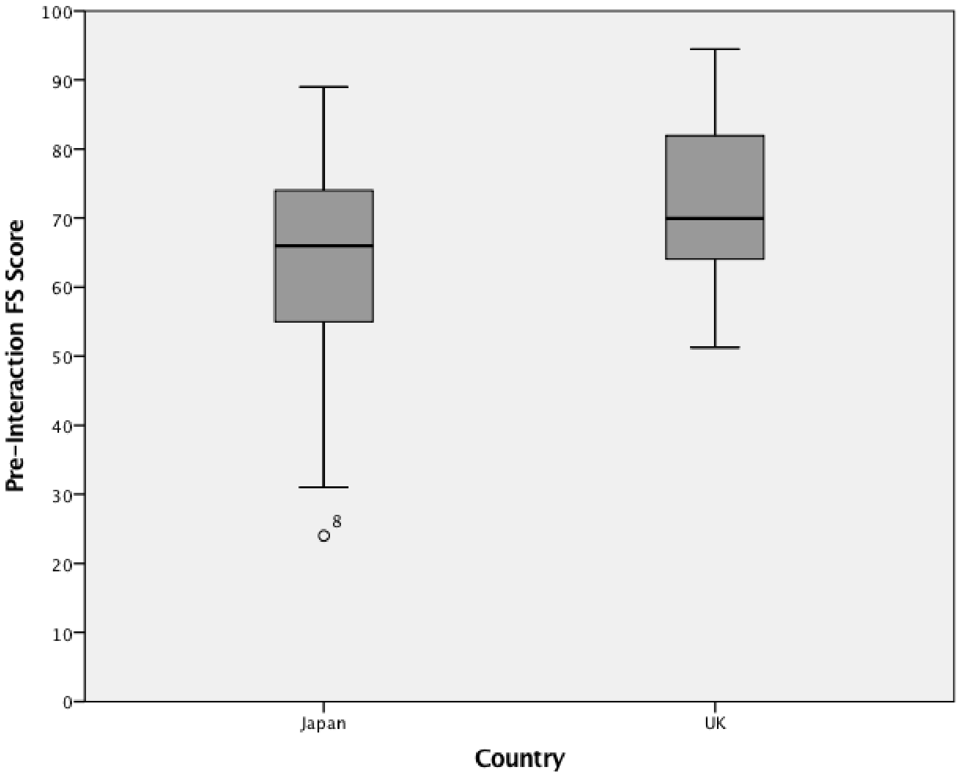


Figure 7.6. Simple boxplot visualising pre-interaction FS scores by country (UK/Japan). There was a significant main effect of country on pre-interaction FS scores.

ures 7.7 and 7.8 for profile plots visualising this for Japan and the UK respectively.

Additional analysis: MIRO and PARO

Given that an interaction period with both MIRO reported here, and PARO reported in Chapter 4, have a significant impact on a participant's FS score a repeated measures ANOVA was also run to compare the interaction effect of robot platform (PARO/MIRO) x FS score delta (as measured

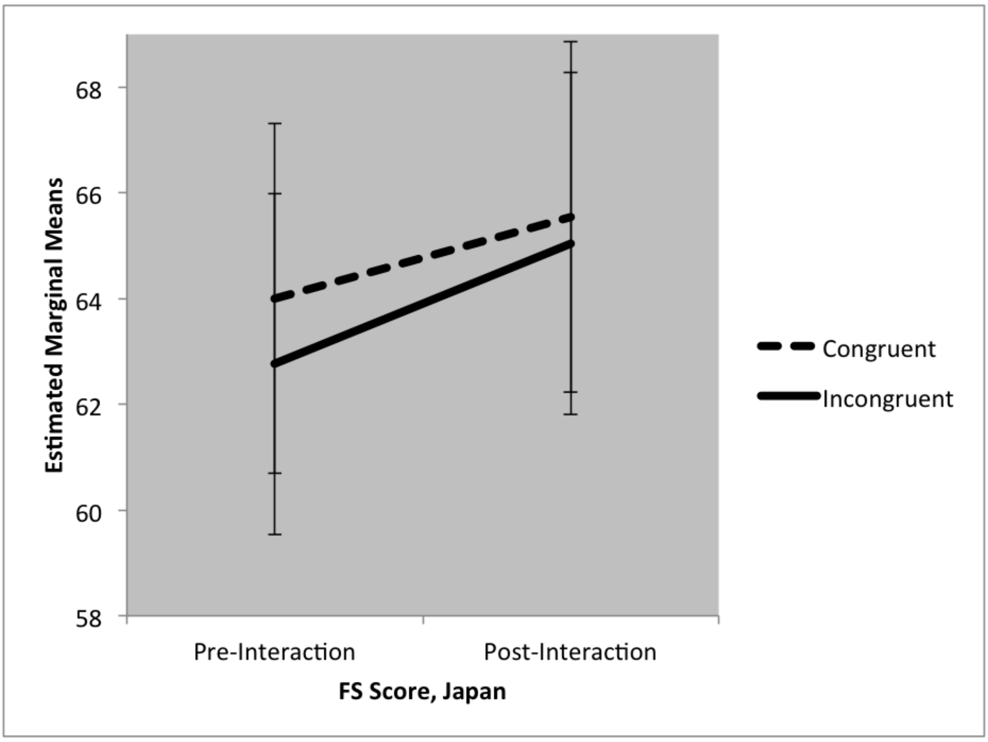


Figure 7.7. Profile plot visualising pre- and post-interaction FS scores by robot condition (congruent/incongruent) for Japan. Plot illustrates significant main effect of FS score from pre- to post-interaction, independent of robot condition. Error bars represent ± 1 Standard Error.

using pre- and post-interaction scores).[7]

There was a significant main effect of FS score, $F(1,\ 154) = 31.81$, $p < .001$, $\eta_p^2 = .02$. This indicates that regardless of which robot platform a participant was interacting with their pre- and post-interaction FS scores significantly differed.

There was a significant robot platform x FS score interaction, $F(1,\ 154) = 6.86$, $p = .010$. This indicates that the difference in pre- and post-

---

[7] This pooling of data was possible as both experiments follow an identical design.

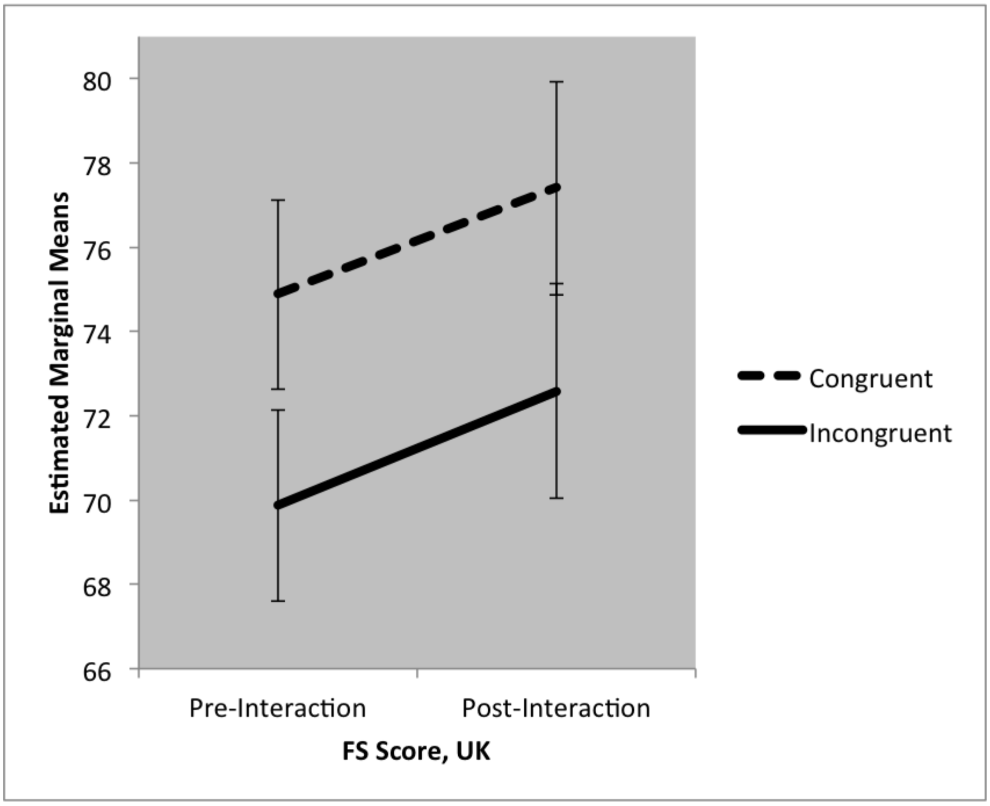


Figure 7.8. Profile plot visualising pre- and post-interaction FS scores by robot condition (congruent/incongruent) for the UK. Plot illustrates significant main effect of FS score from pre- to post-interaction, independent of robot condition. Error bars represent ± 1 Standard Error.

interaction FS scores significantly differed depending upon whether a participant was interacting with PARO or MIRO. See Figure 7.9 for profile plots visualising this for Japan and the UK respectively.

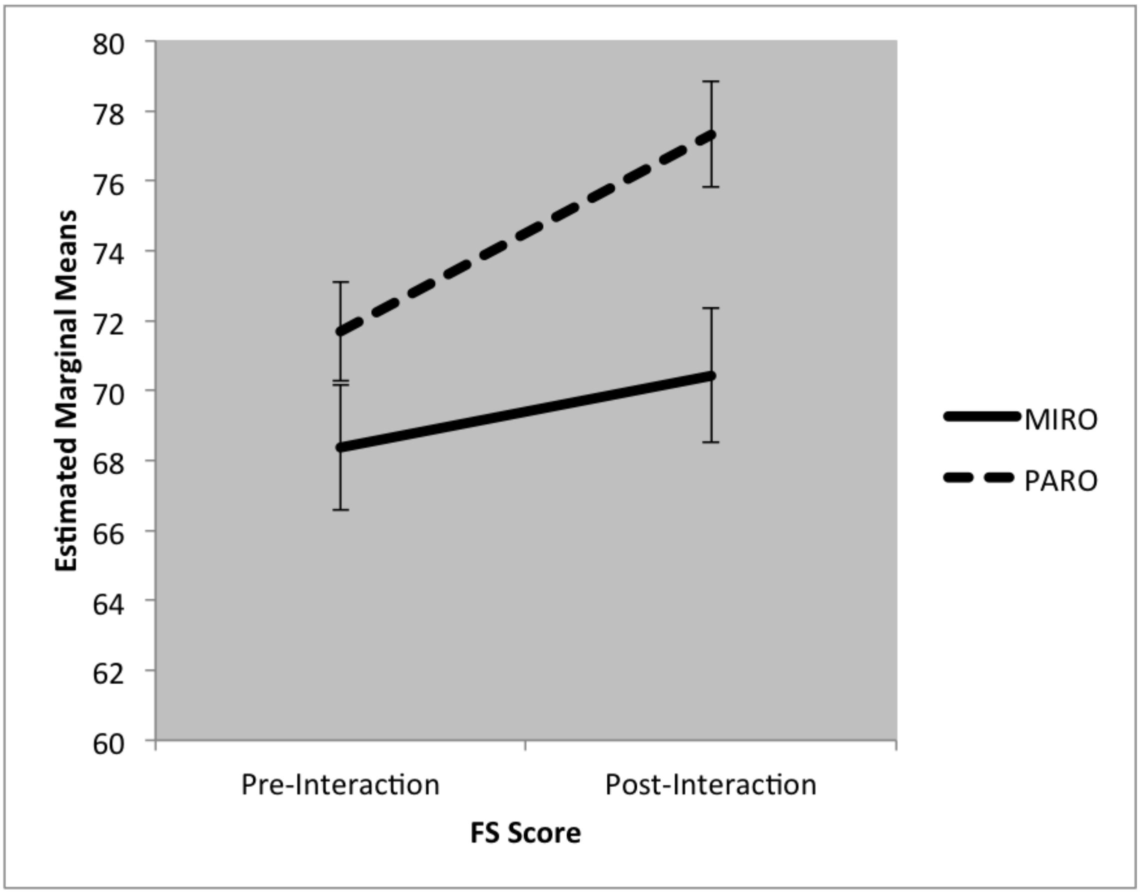


Figure 7.9. Profile plot visualising pre- and post-interaction FS scores by robot platform (MIRO/PARO) for all participants from the UK and Japan. Plot illustrates that the difference in pre- and post-interaction FS scores significantly differed depending upon whether a participant was interacting with PARO or MIRO. Error bars represent ± 1 Standard Error.

7.5 Discussion

The main goal of this chapter was to discover if a positive interactive experience with a social biomimetic robot (as measured via felt security) was related to the behaviour being exhibited by the robot. Would a robot exhibiting 'predictable mammalian behaviour' be more impactful than a robot exhibiting behavioural incongruence? Results show that there is a significant main effect of interacting with MIRO on a change in user felt security as measured from pre- to post-interaction. However, no interaction effect of robot condition on felt security was found. Moreover, this study wanted to address whether there was a cultural aspect to this interaction by comparing a group of participants from the UK with a group from Japan, but no interaction effect of country on felt security was found. Although, a significant main effect of country was found on pre-interaction felt security, with participants from Japan on average reporting lower pre-interaction felt security than those from the UK. However, despite this baseline difference, both Japanese and UK participants had significant positive increases in FS scores post-interaction. Finally, given the results of Chapter 4, which found a significant main effect of PARO on felt security delta, an additional analysis comparing the robot platforms (PARO/MIRO), and their effects on change in felt security from pre- to post-interaction, was also conducted. A significant effect of robot platform on felt security was found, with participants interacting with PARO having greater felt security delta scores than participants interacting with MIRO. Although the effect size for this was small, $\eta_p^2 = .02$, indicating that only 2% of the change in the felt security score could be accounted for by the robot platform being used, it should be noted that these experiments were conducted over a short interaction period in a non-therapeutic context. Were the experiments to be repeated across longer time-spans and with individuals who are explicitly seeking emotional support in a therapeutic capacity a larger effect size might be expected.

The first hypothesis, that spending time interacting with MIRO displaying congruent rather than incongruent behaviour would lead to greater engagement between the user and the robot, and in turn larger felt security delta scores, was not demonstrated by this study. This could be interpreted as a positive outcome, as it shows that the effect of interacting with MIRO may not depend on the precise features of the robot's behaviour. However, as with the study reported in Chapter 4 with the PARO robot, this experiment could have benefitted from a third condition with an inactive robot, or a non-robot control (for example, an interaction period spent manipulating blocks of wood, or even an interaction period spent alone in a room). It had been expected that a difference between the congruent and incongruent versions of MIRO would be observed, especially given the pre-experiment checking phase in which an observable difference between conditions was established. As this was not the outcome it remains unknown what mechanism is effecting the change in participants' felt security scores. Further, it may be that the act of taking part in an experiment is leading to a positive trend in participants' felt security scores. Perhaps on average participants felt more 'secure' (or maybe 'safe') post-interaction than prior to it, as post-interaction their initial nervousness of coming to an unfamiliar room to engage with a novel robot had passed. Therefore future experiments will need to adopt a third control condition, or alternatively a baseline condition needs to be found in which no impact, or a significantly lower impact, on felt security from pre- to post-interaction is being made.

The second hypothesis was that Japanese participants would have a more pronounced FS change score, as measured from pre- to post-interaction, than UK participants. It was expected that lower FS delta scores would be observed in the Japanese group in Condition 2, incongruent, than any other group, based on literature purporting a robot 'craze' in Japan, versus countries like the UK, which may go some way to influencing individuals' mental models of robots, leading to higher expectation from the robot, and as such higher frustration from it when it behaves incongruently. However, this non-significant interaction effect of country could be interpreted as a positive result. As discussed in Chapter 6, recent cross-cultural literature is beginning to report that beliefs regarding the east/west divide are not as straightforward as received opinion would indicate (Bartneck et al., 2007; Haring et al., 2014). Indeed this non-significant interaction effect of country provides insight into how users on an individual level interact with a robot regardless of the everyday social culture their country promotes. Across two countries with reportedly very different socio-cultural relationships with robotics on average (Japan and the UK), pooled participants reported feeling more secure post an interaction with a social biomimetic robot intended for RAT. This result demonstrates the potentially deceptive nature of received opinions regarding cultural difference, and promisingly suggests that the UK market for social biomimetic robots may be no less ready than the Japanese.

7.5.1 Further considerations

The non-significant effect of robot behavioural predictability could have occurred because participants were not noticing the difference between the congruent and incongruent units. Perhaps the behavioural difference between congruent and incongruent conditions was not big enough? However, post-interaction briefings confirmed the findings of the pre-experiment check of detectable difference between conditions. During post-interaction debriefings the overwhelming majority of participants were able to guess which condition they were in once they had been informed of the existence and nature of Conditions 1 and 2. This is indicative of a deal of sympathy being displayed towards the robot by users who were able to enjoy their interactive experience aside from feelings of frustration upon being presented with a behaviourally random unit. Alternatively, it might have been the case that the interaction time was not long enough for participants to become frustrated with an unengaged unit. Perhaps once the novelty effect of engaging with a new robot has passed during a longer term interaction predictability in behaviour becomes an important mechanism of effect driving engagement between a user and their robot?

Further, the randomised behaviour, although apparent, may not have been incongruent enough to warrant a high level of frustration on behalf of the participants. The programme for the incongruent condition was designed to run as congruent but with targets being overridden for each individual orient, thus on occasion the overriding target would match the true congruent target. Perhaps the randomised MIRO unit was simply not random enough? Or, it may have been the case that participants in this experiment were interpreting the otherwise random behaviour of the MIRO as deliberate. An incongruent MIRO may one moment be orienting away from a point of user directed salience, but the next moment, though still acting randomly, directing its attention towards the user by chance, thus creating an illusion of behavioural congruency, whilst simultaneously being rather 'distracted' by other salient items in the environment. One way to address this issue would be to run the experiment again with MIRO in OFF mode, in order to ascertain if levels of felt security remain the same when merely in the presence of a MIRO.

The effect on felt security by robot platform is significantly larger with PARO than with MIRO. This could indicate that there is a feature which differs between these two robots that is acting as a mechanism of effect upon user felt security, but what is that feature? Chapter 5 (Section 5.2) briefly discussed some design aspects which might be influential, such that a positive psychological change is had by a user interacting with a social biomimetic robot. For example, PARO has pettable fur, which MIRO does not. This tactile feature may be important for promoting engagement such that individuals are drawn to the fur more, and in turn feel more relaxed after the act of stroking the robot, which literature on the tactile neurophysiology and related neurochemistry of humans would certainly suggest (as discussed in Chapter 4, Section 4.2.1, see also Barba (1995) for a discussion on how stroking relaxes humans, and Ellingsen et al. (2014) for a discussion on oxytocin release as a result of stroking). Further, there is a deal of difference between PARO's and MIRO's eyes. A robot's eyes are known to be important for encouraging engagement between a robot and a human in a manner akin to their importance in a socialising human-dyad (Bruce et al., 2002). Could it be that how a social biomimetic robot is 'looking' at its user is mechanistic in impacting their felt security? Questions such as these will remain unknown until further systematic research is done on MIRO's design features.

7.5.2 Conclusion

Chapter 7 has provided results that corroborate those reported in Chapter 4. An interaction with a social biomimetic robot was related to an increase in felt security in users. This was despite the marked differences in nationality and culture between participants, and between congruent and incongruent behaviours exhibited by the robot platform MIRO.

As discussed in Chapter 3 (Section 3.2.1), although there are generally applied maxims to be considered during HRI, there are also large differences between individuals in how they function socially, which governs their relationship with agents in their environment. When considering culture as a predictive factor in human-robot interaction outcomes, the impact of individual differences that cut across cultural lines should also be considered. Drawing conclusions about how an individual might respond to a robot based only on where they were born and brought up, without taking individual differences into consideration, may well only serve to hinder the development of, and access to, technology such as robots intended for RAT.

Chapter 8

Conclusions and Future Work

8.1 Summary

In this final chapter a summary of the main contributions of the thesis is provided covering the thesis goals, the theoretical frameworks which inspired the empirical studies, and the outcomes of the empirical studies. The chapter will examine how this body of work has contributed to the field of HRI. It will also consider some of the limitations of the empirical work, how it might be extended in the future, and how it could be applied to the emerging domain of Robot-Assisted Therapy.

8.2 Goals of The Thesis

The goal of this thesis was to present an exploration of the potential for social biomimetic robots in psychological therapeutic interventions, and to advance the field of HRI by using key theories from psychology. To do this the thesis took a comparative approach, proposing that Robot-Assisted Therapy (RAT) could be developed through analogy to Animal-Assisted Therapy (AAT). A framework was developed with which HRI studies can effectively apply existing methodologies from the study of human-other relationships, in order to better understand theory and practice in social robotics. The thesis then applied this framework to the investigation of potential mechanisms of effect of robots developed for RAT, with the intention of positively impacting the well-being of users. Well-being was consistently measured throughout the thesis via changes in the felt security levels of the participants from pre- to post-interaction with a robot. The main contributions of the thesis will here be discussed in turn.

8.3 Theoretical Frameworks

In Chapter 2 the theoretical domain of the thesis was established, by describing the field of AAT and explaining how knowledge of that field can be used to develop RAT. Social needs and well-being were discussed, and the ability for robots to act as animal surrogates in places where animals cannot otherwise go was established. In AAT, use of animals alongside other forms of therapy leads to the support of a client's sense of self and well-being, alleviating loneliness, and contributing to physiological improvements that aid recuperation and healing. This comparative approach framed the empirical work of this thesis, which was an attempt to identify some of the underlying mechanisms of such AAT effects via a comparative methodology within which robots intended for RAT could also be assessed.

Chapter 3 argued that this comparative approach could be achieved by an understanding of how existing methods used to analyse human-other relationships can be transferred to HRI to provide a consistent methodological approach to understanding robots and their impacts on users. This framework places robots alongside other agents with which humans interact, based on their morphology and use, be they humanoid, animal-like, or object-like, establishing conceptual structures used in the remainder of the thesis. Within the context of this broader framework the thesis specifically explored the importance of individual differences, attachment and caregiving styles, the physical engagement levels of a user, and the behavioural predictability of an agent with which the human is interacting.

8.4 Outcomes of Empirical Studies

In Chapter 4 the first experimental study was presented, utilising the social biomimetic robot PARO, which has seal-like morphology, and was designed primarily for use with dementia patients in social facilitation therapy. The mechanisms of effect explored with PARO were the individual differences of the users: their attachment and caregiving styles, as well as their levels of physical engagement with the robot. The work presented in Chapter 4 showed an impact of the robot interaction on a users' felt security, and demonstrated that physical engagement with PARO, which could be rated as 'intimate', resulted in greater levels of post-interaction felt security. This work is the first to demonstrate such an impact on user felt security by an animal-like robot intended for RAT. There was, however, no effect of attachment or caregiving style. Interpreted in terms of the framework presented in Chapter 3 these non-significant effects are important. They demonstrate the need for rigorously exploring measures of interaction, such as attachment style and its influence on human-agent engagement, in order to determine their advantages and limitations in human-non-human interaction research.

The studies using the biomimetic platform MIRO, presented in Chapters 5 and 7, provided further evidence that animal-like robots impact felt security. The non-significant effects of country and behavioural predictability on outcome levels of felt security are consistent with the possibility that these robot-interaction effects on felt security are universal and not restricted by exact behaviour presented by the robot. Taken together with the results from Chapter 4, a pattern can be seen in which robots are being shown to have general applicability.

Ultimately the empirical work presented in the experimental chapters of this thesis has conceived of design for robots intended for RAT with parameters which take into consideration the individual differences of the user, their cultural background, and the features of the robot itself. Via a combination of these factors a clear multi-dimensional understanding of RAT robots can be formed, akin to that developed for selecting AAT co-therapists, which considers individual differences, cultural background, and the features of the animal itself, when pairing the co-therapist with a client. These parameters are not prescriptive, but knowledge of them is useful. This is the first body of work to demonstrate this comparative AAT-RAT approach, and state the potential for such cross-over between disciplines in order to broaden the areas in which such assistive therapies can be given (i.e. no longer restricted to areas in which animals cannot go).

8.5 Additional Contributions

This thesis has also provided a new psychological instrument: the JFSS. A Japanese translation of the English language Felt Security Scale.[1] The translation of this instrument has demonstrated how useful such measures can be for conducting cross-cultural comparative work, if cultural and linguistic differences are taken into consideration during a detailed translation phase.

Finally, testing with the MIRO robot conducted for this thesis has significantly contributed to the development of that platform's non-verbal communication capabilities.

8.6 Limitations and Future Work

Although an interaction period with either PARO or MIRO led to increased felt security across all conditions, the studies in this thesis would benefit from a control condition to establish conclusively that it is the interaction with the robots themselves that leads to the significant increase in post-interaction felt security. As discussed in Chapter 7, a number of other factors may have been contributing to the psychological change. Ap-

---

[1] The limitations of the JFSS are here acknowledged, insofar as further validation testing is required of the instrument on a larger participant sample than that given in Chapter 6.

propriate controls for future studies include a non-robot control, such as a period of time playing a computer game, or interacting with wooden blocks, or an inactive robot control, such as spending a period of time with a robot in OFF mode. The aim of using such controls would be to find a condition in which no impact, or a significantly lower impact, on felt security from pre- to post-interaction is being made. However, although the work presented here did not included a non-robot control condition, a difference was observed between PARO and MIRO on user felt security. An interaction period with MIRO resulted in smaller felt security delta scores than an interaction period with PARO. That MIRO has a smaller impact on felt security than PARO does provide a baseline, of sorts, compared to which PARO provides a stronger effect.

With respect to the study presented in Chapter 7, it must be noted that the interaction period could have been too short to see the potential impact that behavioural incongruency could have on a user. If robots intended for RAT are to be employed similarly to animal AAT co-therapists, their use in longer-term interactions will be expected. In such circumstances, the behavioural repertoire of the robot, and its predictability, is likely to be more important. More generally, longer periods of interaction could have increased psychological benefit (e.g., greater post-interaction increases in felt security) or, conversely, the positive effects might be short-lived. Thus, future studies should look at the longitudinal effects of interacting with social biomimetic robots to establish that increases in, for example, felt security can be had aside from the novelty effects of interacting with an unfamiliar robot.

A further possibility for future work includes the need to develop better, and perhaps automated, instruments for measuring levels of engagement with robots intended for RAT. For instance using computer vision to quantify HRI, programs such as FaceReader to monitor emotional responses, or physiological monitoring (e.g., GSR). The coding work undertaken in Chapter 4 sought to achieve a quantitative, but non-automated, measure of emotional engagement by codifying increasingly intimate physical engagement parameters between a human and a robot. However, the work demonstrated the difficulty in attempting to establish physical engagement parameters that can be consistently used to relate personality (e.g., attachment style) to interaction behaviour.

Given the results from Chapter 4, that a quantitative, but non-automated, measure of emotional engagement coded as *intimate touch* had with PARO lead to greater felt security delta scores, along with the lower post-interaction felt security scores observed after an interaction with furless MIRO as opposed to furry PARO, future work should explore the benefits of direct physical contact, and the mechanisms that mediate them, in more depth. The greater benefit of PARO compared to MIRO is particular intriguing given that work with furless animals, such as turtles, has been suggested to be as effective at reducing anxiety as that with furry animals, such as rabbits (Shiloh et al., 2003). Therefore more research exploring the mechanisms of effect that engage individuals with their AAT co-therapists would also be beneficial for providing designers of RAT robots with feature specifications. This could be achieve by pursuing, for example, research focussed on a better understanding of the C Fibre tactile system in humans and the establishment of consistent physiological measures for assessing contact between a user and a robot intended for RAT. Features which encourage more intimate interactions with users, that could better mimic the interactions observed between humans and animals, would be beneficial to therapists who want to incorporate RAT into their practices. Indeed, one outcome of this thesis, that is of particular importance, is the raised question of the extent to which touching a robot is related to increased felt security. Perhaps robots intended for RAT do need to invite touch, and that is ultimately more important than for them to be behaviourally consistent.

Another issue with this body of work was the use of individuals who are not from an expressly clinical sample. No details were taken from participants as to their current mental health. It is therefore unknown if any participant taking part in the studies presented in this thesis were undertaking any treatment at the time. Working with a clinical sample, who might have lower pre-interaction felt security scores than the participants taking part in the studies presented in this thesis, would provide a stronger demonstration of the potential benefits of robots as therapeutic agents. Further, it would be interesting to extend the analysis presented in Chapter 4 to a participant sample displaying a wider range of attachment and caregiving styles, as the sample in Chapter 4 mostly self-reported as individuals with secure attachment and optimal caregiving styles. In working more closely with clinical populations, and in therapeutic settings, the tailored approach taken by therapists to help clients on a one-to-one level can be better understood, and applied to the development of RAT robots that are accessible to as wide a range of individuals as possible.

Finally, Chapter 4 demonstrated that physical engagement with PARO was not mediated by the attachment and caregiving styles of participants, despite research which outlines that behaviours between humans, at least, can be predicted by such individual differences measures. The chapter further described how attachment theory is currently being applied to human-agent research outside the field of human-human interaction. This is to be encouraged, especially within the context of the comparative framework outlined in Chapter 3. Testing measures that are helpful in assessing the performance of any agents which impact human psychology, in order to explore their advantages as well as their limitations, is crucial to understanding how best to apply them. Given this, the null result of Chapter 4 is of note. Despite predictions derived from an extensive literature, some measures will not transfer well to the new field of HRI, in which a new social entity - the social robot - is being held up against other agents of interaction as part of human-social dyads.

In sum, future studies should focus on the features of the robots themselves, and attempt to understand what robots do to improve their users' psychological well-being whilst understanding the individual differences of users that cut across cultural lines. Future work should also focus on observing interactions between humans and animals, in an attempt to further understand what is being observed during an interaction between a human and an animal-like robot.

8.7 Conclusion

This thesis has shown that by applying theoretical frameworks from the study of human relationships with other humans, animals, or objects, we can advance the research field of human-robot interaction. Moreover, by looking at animals we can extract mechanisms of effect that can be used to design specific features for robots intended for RAT to be used in places where animals cannot otherwise go. A short interaction with a biomimetic robot does lead to an increase in user well-being, as measured by a change in felt security, which is promising for the future development of RAT. Looking ahead, a robot intended for RAT needs to be as flexible as an animal co-therapist, to account for individual differences whilst not prescribing behavioural expectations from a user based on personal or cultural assumptions. An advanced social biomimetic robot that is capable of being tailored to a clinical situation by a therapist in a manner akin to the tailoring of an animal co-therapist in AAT could provide a step-change in the value of robots as therapeutic aids.

Appendix A

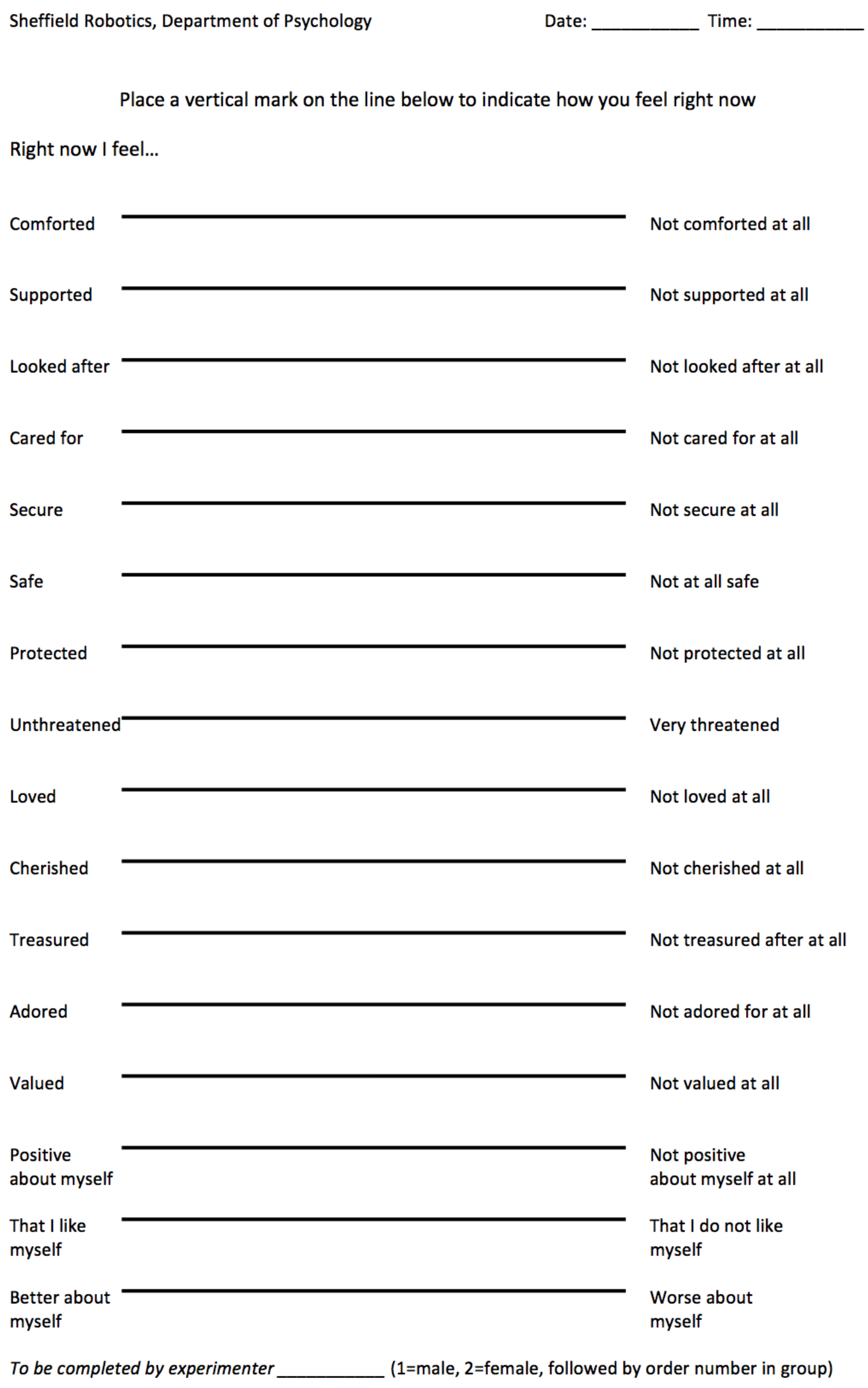
Appendix

A.1 Felt Security Scale (FSS)

Figure A.1 is the 16-item Felt Security Scale (FSS; Luke et al. (2012)) which was taken pre- and post-interaction with both PARO and MIRO for Chapters 4 and 7. It is included here as the version used in this thesis differs to that published by Luke et al. (2012). Rather than being a Likert scale, the FSS used in this thesis was presented as a visual analogue scale (VAS) where each question is answered by making a mark along a 100mm line on either end of which is an item.

A.2 Japanese Felt Security Scale (JFSS)

Figure A.2 is the 11-item Japanese Felt Security Scale, which was taken pre- and post-interaction with MIRO for the Japanese half of the study presented in Chapter 7. The development of the scale is described in Chapter 6. As with the FSS used in this thesis, the JFSS was also presented as a visual analogue scale (VAS) where each question is answered by making a mark along a 100mm line on either end of which is an item-sentence. For English translations of the Japanese see Table 6.2, in Chapter 6.



Sheffield Robotics, Department of Psychology Date: ___________ Time: ___________

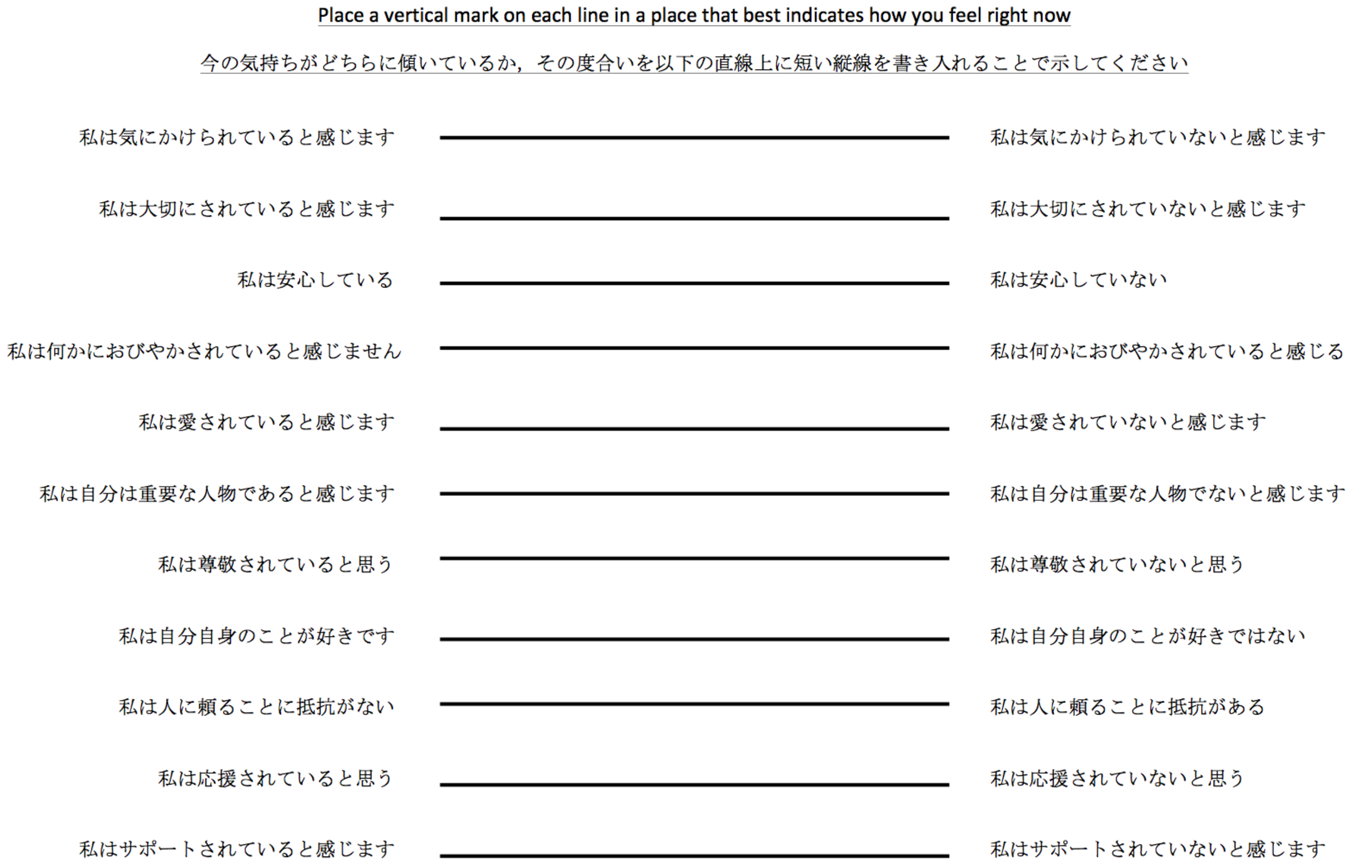
Place a vertical mark on the line below to indicate how you feel right now

Right now I feel…

Comforted	______________________________	Not comforted at all
Supported	______________________________	Not supported at all
Looked after	______________________________	Not looked after at all
Cared for	______________________________	Not cared for at all
Secure	______________________________	Not secure at all
Safe	______________________________	Not at all safe
Protected	______________________________	Not protected at all
Unthreatened	______________________________	Very threatened
Loved	______________________________	Not loved at all
Cherished	______________________________	Not cherished at all
Treasured	______________________________	Not treasured after at all
Adored	______________________________	Not adored for at all
Valued	______________________________	Not valued at all
Positive about myself	______________________________	Not positive about myself at all
That I like myself	______________________________	That I do not like myself
Better about myself	______________________________	Worse about myself

To be completed by experimenter ___________ (1=male, 2=female, followed by order number in group)

Figure A.1. English version of the FSS presented as a VAS.



Place a vertical mark on each line in a place that best indicates how you feel right now

今の気持ちがどちらに傾いているか，その度合いを以下の直線上に短い縦線を書き入れることで示してください

私は気にかけられていると感じます	私は気にかけられていないと感じます
私は大切にされていると感じます	私は大切にされていないと感じます
私は安心している	私は安心していない
私は何かにおびやかされていると感じません	私は何かにおびやかされていると感じる
私は愛されていると感じます	私は愛されていないと感じます
私は自分は重要な人物であると感じます	私は自分は重要な人物でないと感じます
私は尊敬されていると思う	私は尊敬されていないと思う
私は自分自身のことが好きです	私は自分自身のことが好きではない
私は人に頼ることに抵抗がない	私は人に頼ることに抵抗がある
私は応援されていると思う	私は応援されていないと思う
私はサポートされていると感じます	私はサポートされていないと感じます

Figure A.2. Japanese version of the FSS, the JFSS, presented as a VAS.

Bibliography

Ahuvia, A. C. (2005). Beyond the extended self: Loved objects and consumers identity narratives. *Journal of Consumer Research*, 32(1):171–184.

Ainsworth, M. D. S. and Bell, S. M. (1970). Attachment, exploration, and separation: Illustrated by the behavior of one-year-olds in a strange situation. *Child Development*, pages 49–67.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., and Wall, S. N. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Ainsworth, M. D. S. and Parkes, C. M. (1991). Attachments and other affectional bonds across the life cycle. *Attachment Across the Life Cycle*, pages 33–51.

AIST (2006). *Mental commit robot for psychological enrichment.* Retrieved August 30, 2016 from http://www.paro.jp/english/therapy.html.

Allderidge, P. H. (1991). A cat, surpassing in beauty, and other therapeutic animals. *The Psychiatrist*, 15(12):759–762.

Allen, K. M., Blascovich, J., Tomaka, J., and Kelsey, R. M. (1991). Presence of human friends and pet dogs as moderators of autonomic responses to stress in women. *Journal of Personality and Social Psychology*, 61(4):582.

Amichai-Hamburger, Y. and Ben-Artzi, E. (2003). Loneliness and internet use. *Computers in Human Behavior*, 19(1):71–80.

Asimov, I. (1940). Strange playfellow. *Super Science Stories*, pages 67–77.

Ballarini, G. (2003). Pet therapy. animals in human therapy. *Acta Bio Medica Atenei Parmensis*, 74(2):97–100.

Banks, M. R. and Banks, W. A. (2002). The effects of animal-assisted therapy on loneliness in an elderly population in long-term care facilities. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(7):M428–M432.

Banks, M. R., Willoughby, L. M., and Banks, W. A. (2008). Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. *Journal of the American Medical Directors Association*, 9(3):173–177.

Barba, B. E. (1995). The positive influence of animals: Animal-assisted therapy in acute care. *Clinical Nurse Specialist*, 9(4):199–202.

Barker, S. B., Knisely, J. S., Schubert, C. M., Green, J. D., and Ameringer, S. (2015). The effect of an animal-assisted intervention on anxiety and pain in hospitalized children. *Anthrozoös*, 28(1):101–112.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Bartholomew, K. (1990). Avoidance of intimacy: An attachment perspective. *Journal of Social and Personal relationships*, 7(2):147–178.

Bartholomew, K. and Horowitz, L. M. (1991). Attachment styles among young adults: a test of a four-category model. *Journal of Personality and Social Psychology*, 61(2):226.

Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., and Kennsuke, K. (2005). A cross-cultural study on attitudes towards robots. In *HCI International*.

Bartneck, C., Suzuki, T., Kanda, T., and Nomura, T. (2007). The influence of peoples culture and prior experiences with aibo on their attitude towards robots. *AI & Society*, 21(1-2):217–230.

Baxter, P., Kennedy, J., Senft, E., Lemaignan, S., and Belpaeme, T. (2016). From characterising three years of HRI to methodology and reporting recommendations. In *The Eleventh ACM/IEEE International Conference on Human Robot Interation*, pages 391–398. IEEE Press.

Beach, L. R. (1990). *Image theory: Decision making in personal and organizational contexts*. Wiley Chichester.

Beck, A., Hiolle, A., Mazel, A., and Cañamero, L. (2010). Interpretation of emotional body language displayed by robots. In *Proceedings of the 3rd international workshop on Affective Interaction in Natural Environments*, pages 37–42. ACM.

Beck, L. and Madresh, E. A. (2008). Romantic partners and four-legged friends: An extension of attachment theory to relationships with pets. *Anthrozoös*, 21(1):43–56.

Begum, M., Serna, R. W., and Yanco, H. A. (2016). Are robots ready to deliver autism interventions? A comprehensive review. *International Journal of Social Robotics*, pages 1–25.

Belk, R. (1988). *Possessions and self*. Wiley Online Library.

Belpaeme, T., Baxter, P. E., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., Racioppa, S., Kruijff-Korbayová, I., Athanasopoulos, G., Enescu, V., et al. (2012). Multimodal child-robot interaction: Building social bonds. *Journal of Human-Robot Interaction*, 1(2):33–53.

Berget, B. and Braastad, B. O. (2011). Animal-assisted therapy with farm animals for persons with psychiatric disorders. *Annali Dell'Istituto Superiore di Sanità*, 47(4):384–390.

Bharatharaj, J., Huang, L., and Al-Jumaily, A. (2015). Bio-inspired therapeutic pet robots: Review and future direction. In *International Conference on Information and Communication Systems, 2015. IEEE, International Conference*. IEEE.

Bowlby, J. (1969). *Attachment and loss, volume i: Attachment*. Hogarth Press.

Braun, C., Stangler, T., Narveson, J., and Pettingell, S. (2009). Animal-assisted therapy as a pain relief intervention for children. *Complementary Therapies in Clinical Practice*, 15(2):105–109.

Braungart, J. M. and Stifter, C. A. (1991). Regulation of negative reactivity during the strange situation: Temperament and attachment in 12-month-old infants. *Infant Behavior and Development*, 14(3):349–364.

Breazeal, C. and Scassellati, B. (1999). How to build robots that make friends and influence people. In *Intelligent Robots and Systems. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference*, volume 2, pages 858–863. IEEE.

Brennan, K., Clark, C., and Shaver, P. (1998). Self-report measurement of adult attachment: An integrative overview. In *Attachment theory and close relationships*, pages 46–76. Guilford Press, New York.

Brickel, C. M. (1985). Initiation and maintenance of the human-animal bond: Familial roles from a learning perspective. *Marriage & Family Review*, 8(3-4):31–48.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3):185–216.

Broekens, J., Heerink, M., and Rosendal, H. (2009). Assistive social robots in elderly care: a review. *Gerontechnology*, 8(2):94–103.

Bruce, A., Nourbakhsh, I., and Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. In *Robotics and Automation. Proceedings. ICRA'02. IEEE International Conference*, volume 4, pages 4138–4142. IEEE.

Bucy, E. P. and Newhagen, J. E. (2004). *Media access: Social and psychological dimensions of new technology use.* Psychology Press.

Burke, S. (1992). In the presence of animals. Health professionals no longer scoff at the therapeutic effects of pets. *US News & World Report*, 112(7):64.

Burrows, K. E., Adams, C. L., and Spiers, J. (2008). Sentinels of safety: Service dogs ensure safety and enhance freedom and well-being for families with autistic children. *Qualitative Health Research*, 18(12):1642–1649.

Bustad, L. K. (1981). *Animals, aging, and the aged.* University of Minnesota Press.

Carpenter, J. (2016). *Culture and human-robot interaction in militarized spaces: A war story.* Ashgate Publishing, Ltd.

Carr, S. and Rockett, B. (2014). Animals and attachment theory. *Society and Animals*, 22(4):415–433.

Carroll, J. (2016). *Top robotics companies for 2016 named by Robotics Business Review.* Retrieved August 30, 2016 from http://www.vision-systems.com/articles/2016/03/top-robotics-companies-for-2016-named-by-robotics-business-review.html.

Caspi, A. and Shiner, R. L. (2006). Personality development. *Handbook of Child Psychology.*

Chen, Z.-H., Deng, Y.-C., Chou, C.-Y., and Chan, T.-W. (2005). Motivating learners by nurturing animal companions: My-pet and our-pet. In *AIED*, pages 136–143.

Chugani, H. T., Behen, M. E., Muzik, O., Juhász, C., Nagy, F., and Chugani, D. C. (2001). Local brain functional activity following early deprivation: a study of postinstitutionalized romanian orphans. *Neuroimage*, 14(6):1290–1301.

Clarke-Stewart, K. A. (1989). Infant day care: Maligned or malignant? *American Psychologist*, 44(2):266.

Clarkson, E. and Arkin, R. C. (2007). Applying heuristic evaluation to human-robot interaction systems. In *Flairs Conference*, pages 44–49.

Colby, P. M. and Sherman, A. (2002). Attachment styles impact on pet visitation effectiveness. *Anthrozoös*, 15(2):150–165.

Collins, E. C., Millings, A., and Prescott, T. J. (2013). Attachment to assistive technology: A new conceptualisation. In *Proceedings of the 12th European AAATE Conference (Association for the Advancement of Assistive Technology in Europe).*

Collins, E. C., Prescott, T. J., Mitchinson, B., and Conran, S. (2015). Miro: A versatile biomimetic edutainment robot. In *Advances in Computer Entertainment.* Springer.

Collins, N. L. and Feeney, B. C. (2000). A safe haven: An attachment theory perspective on support seeking and caregiving in intimate relationships. *Journal of Personality and Social Psychology*, 78(6):1053.

Collis, G. and McNicholas, J. (1998). A theoretical basis for health benefits of pet ownership. *Companion Animals in Human Health*, pages 105–22.

Connor, K. and Miller, J. (2000). Animal-assisted therapy: An in-depth look. *Dimensions of Critical Care Nursing*, 19(3):20–hyhen.

Cornwell, E. Y. and Waite, L. J. (2009). Social disconnectedness, perceived isolation, and health among older adults. *Journal of Health and Social Behavior*, 50(1):31–48.

Courtney, A. J. (1986). Chinese population stereotypes: Color associations. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 28(1):97–99.

Crawford, E. K., Worsham, N. L., and Swinehart, E. R. (2006). Benefits derived from companion animals, and the use of the term attachment. *Anthrozoös*, 19(2):98–112.

Crockenberg, S. B. (1981). Infant irritability, mother responsiveness, and social support influences on the security of infant-mother attachment. *Child Development*, pages 857–865.

Darcy, K., Davila, J., and Beck, J. G. (2005). Is social anxiety associated with both interpersonal avoidance and interpersonal dependence? *Cognitive Therapy and Research*, 29(2):171–186.

Darwin, C. (2002). *The expression of the emotions in man and animals.* Oxford University Press.

Dautenhahn, K. (2007). Methodology and themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems.*

Dautenhahn, K. and Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, 12(1):1–35.

Davis, M. H., Association, A. P., et al. (1980). *A multidimensional approach to individual differences in empathy.* American Psychological Association.

Davis, N., Custer, K., Marcey, M., and Solarz, V. (1996). *Once Upon a Time – : Therapeutic Stories that Teach & Heal.* N. Davis.

Dean, P., Redgrave, P., and Westby, G. (1989). Event or emergency? Two response systems in the mammalian superior colliculus. *Trends in Neurosciences*, 12(4):137–147.

Denzau, A. T. and North, D. C. (1994). Shared mental models: Ideologies and institutions. *Kyklos*, 47(1):3–31.

Diehl, J. J., Schmitt, L. M., Villano, M., and Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in Autism Spectrum Disorders*, 6(1):249–262.

Diener, E. and Emmons, R. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47:1105–1117.

Diesendruck, G. and Perez, R. (2015). Toys are me: Childrens extension of self to objects. *Cognition*, 134:11–20.

Dunbar, R. I. and Shultz, S. (2007). Evolution in the social brain. *Science*, 317(5843):1344–1347.

Edgar, J., Nicol, C., Clark, C., and Paul, E. (2012). Measuring empathic responses in animals. *Applied Animal Behaviour Science*, 138(3):182–193.

Egeland, B. and Farber, E. A. (1984). Infant-mother attachment: Factors related to its development and changes over time. *Child Development*, pages 753–771.

Ellingsen, D.-M., Wessberg, J., Chelnokova, O., Olausson, H., Laeng, B., and Leknes, S. (2014). In touch with your emotions: Oxytocin and touch change social impressions while others facial expressions can alter touch. *Psychoneuroendocrinology*, 39:11–20.

Emanuel, R. and Weir, S. (1976). Catalysing communication in an autistic child in a logo-like learning environment. In *AISB (ECAI)*, pages 118–129.

Epley, N., Waytz, A., and Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4):864.

Fensom, A. (2015). *Japan's robot revolution*. Retrieved August 30, 2016 from http://thediplomat.com/2015/07/japans-robot-revolution/.

Filan, S. L. and Llewellyn-Jones, R. H. (2006). Animal-assisted therapy for dementia: A review of the literature. *International Psychogeriatrics*, 18(4):597–612.

Fine, A. H. (2006). Incorporating animal-assisted therapy into psychotherapy: Guidelines and suggestions for therapists. *Handbook on Animal-Assisted Therapy: Theoretical Foundations and Guidelines for Practice*, pages 167–206.

Fine, A. H. (2015). *Handbook on animal-assisted therapy: Theoretical foundations and guidelines for practice*. Academic Press.

Fine, A. H., Dennis, A. L., and Bowers, C. (2011). Incorporating animal-assisted interventions in therapy with boys at risk. *Engaging Boys in Treatment: Creative Approaches to the Therapy-Process*, pages 115–134.

Fink, J., Mubin, O., Kaplan, F., and Dillenbourg, P. (2012). Anthropomorphic language in online forums about roomba, AIBO and the iPad. In *Advanced Robotics and its Social Impacts (ARSO), 2012 IEEE Workshop*, pages 54–59. IEEE.

Fowler, J. H. and Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *BMJ*, 337:a2338.

Fraley, R. C. and Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In Rholes, J. A. S. . W. S., editor, *Attachment theory and close relationships*, pages 77–114. Guilford Press, New York.

Friedman, B., Kahn Jr, P. H., and Hagman, J. (2003). Hardware companions?: What online aibo discussion forums reveal about the human-robotic relationship. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 273–280. ACM.

Friedmann, E., Katcher, A. H., Lynch, J. J., and Thomas, S. A. (1980). Animal companions and one-year survival of patients after discharge from a coronary care unit. *Public Health Reports*, 95(4):307.

Friedmann, E., Son, H., and Tsai, C. (2000). The animal-human bond: Health and wellness. *Handbook on Animal-Assisted Therapy: Theoretical Foundations and Guidelines for Practice*, pages 41–58.

Gentner, D. and Stevens, A. L. (2014). *Mental models*. Psychology Press.

Gillath, O., Hart, J., Noftle, E. E., and Stockdale, G. D. (2009). Development and validation of a state adult attachment measure (SAAM). *Journal of Research in Personality*, 43(3):362–373.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, 7(1):7–28.

Grice, H. P., Cole, P., and Morgan, J. L. (1975). Syntax and semantics. *Logic and Conversation*, 3:41–58.

Guerrero, L. K. (1996). Attachment-style differences in intimacy and involvement: A test of the four-category model. *Communications Monographs*, 63(4):269–292.

Gulrez, T., Nefti-Meziani, S., McEvoy, P., and Hodgson, A. (2016). Loneliness kills: can autonomous systems and robotics assist in providing solutions? *International Journal of Swarm Intelligence and Evolutionary Computation*, 5(1):e113.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001). A computational model of action selection in the basal ganglia. A new functional anatomy. *Biological Cybernetics*, 84(6):401–410.

Hall, E. T. (1966). *The hidden dimension.* Doubleday & Co.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold Jr, A. R., Durbin, M., Edmonson, M. S., Fischer, J., Hymes, D., Kimball, S. T., et al. (1968). Proxemics [and comments and replies]. *Current Anthropology*, pages 83–108.

Haring, K. S., Silvera-Tawil, D., Matsumoto, Y., Velonaki, M., and Watanabe, K. (2014). Perception of an android robot in Japan and Australia: A cross-cultural comparison. In *International Conference on Social Robotics*, pages 166–175. Springer.

Haring, M., Bee, N., and André, E. (2011). Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *RO-MAN, 2011 IEEE*, pages 204–209. IEEE.

Harlow, H. F. and Harlow, M. K. (1969). Effects of various mother-infant relationships on rhesus monkey behaviors. *Determinants of Infant Behavior*, 4:15–36.

Hazan, C. and Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, 52(3):511.

Hazan, C. and Zeifman, D. (1999). *Pair bonds as attachments: Evaluating the evidence.* Guilford Press.

Heerink, M., Albo-Canals, J., Valenti-Soler, M., Martinez-Martin, P., Zondag, J., Smits, C., and Anisuzzaman, S. (2013). Exploring requirements and alternative pet robots for robot assisted therapy with older adults with dementia. In *International Conference on Social Robotics*, pages 104–115. Springer.

Herdman, M., Fox-Rushby, J., and Badia, X. (1997). ‘Equivalence’ and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research*, 6(3):0–0.

Hiroi, Y. and Ito, A. (2008). Are bigger robots scary? The relationship between robot size and psychological threat. In *2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 546–551. IEEE.

Holmén, K., Ericsson, K., and Winblad, B. (1999). Quality of life among the elderly. *Scandinavian Journal of Caring Sciences*, 13(2):91–95.

Holmes, J. G. and Rempel, J. K. (1989). Trust in close relationships. In Hendrick, C., editor, *Close relationships. Review of Personality and Social Psychology*, pages 187–220. Sage Publications, Inc, Thousand Oaks, CA, US.

Huang, L. and Gillan, D. (2014). An exploration of robot builders’ emotional responses to their tournament robots. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 2013–2017. SAGE Publications.

Hulin, C. L. (1987). A psychometric theory of evaluations of item and scale translations fidelity across languages. *Journal of Cross-Cultural Psychology*, 18(2):115–142.

Hurlbert, A. C. and Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, 17(16):R623–R625.

Jenkins, R., Rose, J., and Lovell, C. (1997). Psychological well-being of staff working with people who have challenging behaviour. *Journal of Intellectual Disability Research*, 41(6):502–511.

Jipson, J. L., Gülgöz, S., and Gelman, S. A. (2016). Parent–child conversations regarding the ontological status of a robotic dog. *Cognitive Development*, 39:21–35.

Jones, T., Lawson, S., and Mills, D. (2008). Interaction with a zoomorphic robot that exhibits canid mechanisms of behaviour. In *Robotics and Automation. ICRA 2008. IEEE International Conference*, pages 2128–2133. IEEE.

Kaler, S. R. and Freeman, B. (1994). Analysis of environmental deprivation: Cognitive and social development in romanian orphans. *Journal of Child Psychology and Psychiatry*, 35(4):769–781.

Kang, M. (2011). *Sublime dreams of living machines.* Harvard University Press.

Kaplan, F. (2004). Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International Journal of Humanoid Robotics*, 1(03):465–480.

Kazdin, A. E. and Fine, A. (2010). Methodological standards and strategies for establishing the evidence base of animal-assisted therapies. *Handbook on Animal-Assisted Therapy.*

Keefer, L. A., Landau, M. J., Rothschild, Z. K., and Sullivan, D. (2012). Attachment to objects as compensation for close others' perceived unreliability. *Journal of Experimental Social Psychology*, 48(4):912–917.

Kerepesi, A., Kubinyi, E., Jonsson, G., Magnusson, M., and Miklosi, A. (2006). Behavioural comparison of human–animal (dog) and human–robot (AIBO) interactions. *Behavioural Processes*, 73(1):92–99.

Kidd, C. D., Taggart, W., and Turkle, S. (2006). A sociable robot to encourage social interaction among the elderly. In *Robotics and Automation: ICRA 2006. Proceedings 2006 IEEE International Conference*, pages 3972–3976. IEEE.

Kiesler, T. and Kiesler, S. (2004). My pet rock and me: An experimental exploration of the self extension concept. *Human-Computer Interaction Institute*, page 85.

Klein, B., Gaedt, L., and Cook, G. (2013). Emotional robots: Principles and experiences with paro in denmark, germany, and the uk. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, 26(2):89.

Kogut, T. and Kogut, E. (2011). Possession attachment: Individual differences in the endowment effect. *Journal of Behavioral Decision Making*, 24(4):377–393.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., and Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53(9):1017.

Kruger, K. A., Serpell, J. A., and Fine, A. (2006). Animal-assisted interventions in mental health: Definitions and theoretical foundations. *Handbook on Animal-Assisted Therapy: Theoretical Foundations and Guidelines for Practice*, 2:21–38.

Kurdek, L. A. (2008). Pet dogs as attachment figures. *Journal of Social and Personal Relationships*, 25(2):247–266.

Kuss, D, J., Griffiths, M, D., Karila, L., and Billieux, J. (2014). Internet addiction: A systematic review of epidemiological research for the last decade. *Current Pharmaceutical Design*, 20(25):4026–4052.

Kutner, J. S., Smith, M. C., Corbin, L., Hemphill, L., Benton, K., Mellis, B. K., Beaty, B., Felton, S., Yamashita, T. E., Bryant, L. L., et al. (2008). Massage therapy versus simple touch to improve pain and mood in patients with advanced cancer: A randomized trial. *Annals of Internal Medicine*, 149(6):369–379.

Lambert, D. (2004). *Body language*. HarperCollins.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, pages 159–174.

Lazar, J., Jones, A., and Shneiderman, B. (2006). Workplace user frustration with computers: An exploratory investigation of the causes and severity. *Behaviour & Information Technology*, 25(03):239–251.

Leiber, J. (1997). Natures experiments, society's closures. *Journal for the Theory of Social Behaviour*, 27(2-3):325–343.

Leslie, A. M. (1994). Tomm, toby, and agency: Core architecture and domain specificity. In Hirschfeld, L. A. and Gelman, S. A., editors, *Mapping the mind*, pages 119–148. Cambridge University Press. Cambridge Books Online.

Libin, A. and Cohen-Mansfield, J. (2004). Therapeutic robocat for nursing home residents with dementia: Preliminary inquiry. *American Journal of Alzheimer's Disease and Other Dementias*, 19(2):111–116.

Löken, L. S. and Olausson, H. (2010). The skin as a social organ. *Experimental Brain Research*, 204(3):305–314.

Luke, M. A., Sedikides, C., and Carnelley, K. (2012). Your love lifts me higher! The energizing quality of secure relationships. *Personality and Social Psychology Bulletin*, 38(6):721–733.

MacDorman, K. F. and Ishiguro, H. (2006). Opening pandoras uncanny box: Reply to commentaries on 'the uncanny advantage of using androids in social and cognitive science research'. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 7(3):361–368.

MacDorman, K. F., Vasudevan, S. K., and Ho, C.-C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society*, 23(4):485–510.

MacNamara, M., Moga, J., and Pachel, C. (2015). Whats love got to do with it? Selecting animals for animal-assisted mental health interventions. *Handbook on Animal-Assisted Therapy.*

Magai, C., Cohen, C., Gomberg, D., Malatesta, C., and Culver, C. (1996). Emotional expression during mid-to late-stage dementia. *International Psychogeriatrics*, 8(03):383–395.

Maier, M. A., Barchfeld, P., Elliot, A. J., and Pekrun, R. (2009). Context specificity of implicit preferences: The case of human preference for red. *Emotion*, 9(5):734.

Mallinckrodt, B. and Wang, C.-C. (2004). Quantitative methods for verifying semantic equivalence of translated research instruments: A Chinese version of the Experiences in Close Relationships Scale. *Journal of Counseling Psychology*, 51(3):368.

Manav, B. (2007). Color-emotion associations and color preferences: A case study for residences. *Color Research & Application*, 32(2):144–150.

Mantovani, G. (1996). Social context in HCl: A new framework for mental models, cooperation, and communication. *Cognitive Science*, 20(2):237–269.

Markus, H. R. and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2):224.

Marvin, R. S. and Britner, P. A. (1999). *Normative development: The ontogeny of attachment.* Guilford Press.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4):370.

Maslow, A. H., Frager, R., Fadiman, J., McReynolds, C., and Cox, R. (1970). *Motivation and personality*, volume 2. Harper & Row New York.

Mäthger, L. M. and Hanlon, R. T. (2007). Malleable skin coloration in cephalopods: Selective reflectance, transmission and absorbance of light by chromatophores and iridophores. *Cell and Tissue Research*, 329(1):179–186.

McGlone, F., Olausson, H., Boyle, J., Jones-Gotman, M., Dancer, C., Guest, S., and Essick, G. (2012). Touching and feeling: Differences in pleasant touch processing between glabrous and hairy skin in humans. *European Journal of Neuroscience*, 35(11):1782–1788.

McGlone, F., Vallbo, A. B., Olausson, H., Loken, L., and Wessberg, J. (2007). Discriminative touch and emotional touch. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 61(3):173.

Mendl, M., Burman, O. H., and Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696):2895–2904.

Meyer, A. S., Sleiderink, A. M., and Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, 66(2):B25–B33.

Miklósi, Á. and Gácsi, M. (2012). On the utilisation of social animals as a model for social robotics. *Frontiers in Psychology*, 3(75).

Mikulincer, M. (1986). Shaver, p. r.(2007). *Attachment in adulthood: Structure, dynamics, and change.*

Miller, J. and Ingram, L. (2000). Perioperative nursing and animal-assisted therapy. *AORN Journal*, 72(3):477–483.

Millings, A. and Collins, E. C. (2016). Human relationships with living machines. In Prescott, T. and Verschure, P., editors, *Living Machines: A Handbook of Research in Biomimetic and Biohybrid Systems*. Oxford University Press. In Press.

Millings, A., Walsh, J., Hepper, E., and OBrien, M. (2012). Good partner, good parent responsiveness mediates the link between romantic attachment and parenting style. *Personality and Social Psychology Bulletin*, page 0146167212468333.

Morahan-Martin, J. and Schumacher, P. (2003). Loneliness and social uses of the internet. *Computers in Human Behavior*, 19(6):659–671.

Mordoch, E., Osterreicher, A., Guse, L., Roger, K., and Thompson, G. (2013). Use of social commitment robots in the care of elderly people with dementia: A literature review. *Maturitas*, 74(1):14–20.

Mossello, E., Ridolfi, A., Mello, A. M., Lorenzini, G., Mugnai, F., Piccini, C., Barone, D., Peruzzi, A., Masotti, G., and Marchionni, N. (2011). Animal-assisted activity and emotional status of patients with alzheimer's disease in day care. *International Psychogeriatrics*, 23(06):899–905.

Murray, S. L., Holmes, J. G., and Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology*, 78(3):478.

Murray, S. L., Rose, P., Holmes, J. G., Derrick, J., Podchaski, E. J., Bellavia, G., and Griffin, D. W. (2005). Putting the partner within reach: A dyadic perspective on felt security in close relationships. *Journal of Personality and Social Psychology*, 88(2):327.

Murthy, R., Bearman, G., Brown, S., Bryant, K., Chinn, R., Hewlett, A., George, B. G., Goldstein, E. J., Holzmann-Pazgal, G., Rupp, M. E., et al. (2015). Animals in healthcare facilities: Recommendations to minimize potential risks. *Infection Control & Hospital Epidemiology*, 36(05):495–516.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., and Hagita, N. (2009). Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proceedings of the 4th ACM/IEEE International Conference on Human robot interaction*, pages 61–68. ACM.

Nakao, T. and Kato, K. (2004). Developing the Japanese version of the Adult Attachment Style Scale (ECR). *Shinrigaku Kenkyu: The Japanese Journal of Psychology*, 75(2):154–159.

Nass, C. and Reeves, B. (1996). *The media equation: How people treat computers, televisions, and new media as real people and places*. Cambridge University Press.

Nelson, D. L. and Quick, J. C. (1991). Social support and newcomer adjustment in organizations: Attachment theory at work? *Journal of Organizational Behavior*, 12(6):543–554.

Nightingale, F. (1969). *Notes on Nursing: What It Is, and What It If Not.* Dover Publications, New York.

Nimer, J. and Lundahl, B. (2007). Animal-assisted therapy: A meta-analysis. *Anthrozoös*, 20(3):225–238.

O'Haire, M. E. (2013). Animal-assisted intervention for autism spectrum disorder: A systematic literature review. *Journal of Autism and Developmental Disorders*, 43(7):1606–1622.

Olausson, H., Wessberg, J., McGlone, F., and Vallbo, Å. (2010). The neurophysiology of unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews*, 34(2):185–191.

Owens, E. W., Behun, R. J., Manning, J. C., and Reid, R. C. (2012). The impact of internet pornography on adolescents: A review of the research. *Sexual Addiction & Compulsivity*, 19(1-2):99–122.

Panzer-Koplow, S. L. (2000). *Effects of animal-assisted therapy on depression and morale among nursing home residents.* Rutgers.

Parish-Plass, N. (2008). Animal-assisted therapy with children suffering from insecure attachment due to abuse and neglect: A method to lower the risk of intergenerational transmission of abuse? *Clinical Child Psychology and Psychiatry*, 13(1):7–30.

Parkin, D. J. (1999). Mementoes as transitional objects in human displacement. *Journal of Material Culture*, 4(3):303–320.

Parshall, D. P. (2003). Research and reflection: Animal-assisted therapy in mental health settings. *Counseling and Values*, 48(1):47–56.

Payne, E., Bennett, P. C., and McGreevy, P. D. (2015). Current perspectives on attachment and bonding in the dog–human dyad. *Psychology Research and Behavior Management*, 8:71.

Pearson, M. J., Mitchinson, B., Sullivan, J. C., Pipe, A. G., and Prescott, T. J. (2011). Biomimetic vibrissal sensing for robots. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366(1581):3085–3096.

Pearson, M. J., Pipe, A. G., Melhuish, C., Mitchinson, B., and Prescott, T. J. (2007). Whiskerbot: A robotic active touch system modeled on the rat whisker sensory system. *Adaptive Behavior*, 15(3):223–240.

Penninx, B. W., Van Tilburg, T., Kriegsman, D. M., Deeg, D. J., Boeke, A. J. P., and van Eijk, J. T. M. (1997). Effects of social support and personal coping resources on mortality in older age: The longitudinal aging study amsterdam. *American Journal of Epidemiology*, 146(6):510–519.

Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(03):715–734.

Prescott, T. J., Redgrave, P., and Gurney, K. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behavior*, 7(1):99–127.

Prodhan, G. (2015). *China to have most robots in world by 2017*. Retrieved February 1, 2015 from http://www.reuters.com/article/robots-china-idusl6n0vf52o20150205.

Prothmann, A., Ettrich, C., and Prothmann, S. (2009). Preference for, and responsiveness to, people, dogs and objects in children with autism. *Anthrozoös*, 22(2):161–171.

Reeves, B. and Nass, C. (1996). *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press.

Riek, L. D. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1).

Robertson, J. (2007). Robo sapiens Japanicus: Humanoid robots and the posthuman family. *Critical Asian Studies*, 39(3):369–398.

Robins, B., Amirabdollahian, F., and Dautenhahn, K. (2013). Investigating child-robot tactile interactions: A taxonomical classification of tactile behaviour of children with autism towards a humanoid robot. In *The Sixth International Conference on Advances in Computer-Human Interactions (ACHI)*, pages 89–94. Citeseer.

Robins, B., Dautenhahn, K., Te Boekhorst, R., and Billard, A. (2005). Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society*, 4(2):105–120.

Robins, R. W., Caspi, A., and Moffitt, T. E. (2000). Two personalities, one relationship: Both partners' personality traits shape the quality of their relationship. *Journal of Personality and Social Psychology*, 79(2):251.

Robins, R. W., Caspi, A., and Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality*, 70(6):925–964.

Robinson, H., MacDonald, B., Kerse, N., and Broadbent, E. (2013). The psychosocial effects of a companion robot: A randomized controlled trial. *Journal of the American Medical Directors Association*, 14(9):661–667.

Rodogno, R. (2015). Social robots, fiction, and sentimentality. *Ethics and Information Technology*, pages 1–12.

Rosenthal-von der Pütten, A. M., Krämer, N. C., Hoffmann, L., Sobieraj, S., and Eimler, S. C. (2013). An experimental study on emotional reactions towards a robot. *International Journal of Social Robotics*, 5(1):17–34.

Rothbaum, F., Weisz, J., Pott, M., Miyake, K., and Morelli, G. (2000). Attachment and culture: Security in the United States and Japan. *American Psychologist*, 55(10):1093.

Rushforth, J. (2009). Can attachment security priming influence experiences of a social task in a socially anxious group. *Unpublished doctoral dissertation. University of Southampton, England.*

Šabanović, S. (2010). Robots in society, society in robots. *International Journal of Social Robotics*, 2(4):439–450.

Šabanović, S., Bennett, C. C., Chang, W., and Huber, L. (2013). PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference*, pages 1–6. IEEE.

Šabanović, S., Bennett, C. C., and Lee, H. R. (2014). Towards culturally robust robots: A critical social perspective on robotics and culture. In *Proc. HRI Workshop on Culture-Aware Robotics 2014*.

Samani, H. A., Cheok, A. D., Tharakan, M. J., et al. (2011). A design process for lovotics. Human-robot personal relationships. *Human-Robot Personal Relationships*, pages 118–125.

Sardar, A., Joosse, M., Weiss, A., and Evers, V. (2012). Don't stand so close to me: Users' attitudinal and behavioral responses to personal space invasion by robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 229–230. ACM.

Sarwar, M. and Soomro, T. R. (2013). Impact of smartphones on society. *European Journal of Scientific Research*, 98(2):216–226.

Scassellati, B., Admoni, H., and Mataric, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294.

Serpell, J. A. (2006). Animal-assisted interventions in historical perspective. *Handbook on Animal-Assisted Therapy: Theoretical Foundations and Guidelines for Practice*, 2:3–20.

Sharkey, N. and Sharkey, A. (2010). The crying shame of robot nannies: An ethical appraisal. *Interaction Studies*, 11(2):161–190.

Sharpe, T. M., Killen, J. D., Bryson, S. W., Shisslak, C. M., Estes, L. S., Gray, N., Crago, M., and Taylor, C. B. (1998). Attachment style and weight concerns in preadolescent and adolescent girls. *International Journal of Eating Disorders*, 23(1):39–44.

Shaver, P. R., Mikulincer, M., and Shemesh-Iron, M. (2010). A behavioral systems perspective on prosocial behavior. *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature*, pages 73–92.

Shibata, T. (2016). *PARO.* Retrieved February 2, 2016 from http://www.parorobots.com/.

Shibata, T., Mitsui, T., Wada, K., Touda, A., Kumasaka, T., Tagami, K., and Tanie, K. (2001). Mental commit robot and its application to therapy of children. In *Advanced Intelligent Mechatronics. Proceedings. 2001 IEEE/ASME International Conference*, volume 2, pages 1053–1058. IEEE.

Shiloh, S., Sorek, G., and Terkel, J. (2003). Reduction of state-anxiety by petting animals in a controlled laboratory experiment. *Anxiety, Stress, and Coping*, 16(4):387–395.

Souter, M. A. and Miller, M. D. (2007). Do animal-assisted activities effectively treat depression? A meta-analysis. *Anthrozoös*, 20(2):167–180.

Sparrow, R. and Sparrow, L. (2006). In the hands of machines? The future of aged care. *Minds and Machines*, 16(2):141–161.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., and Vagg, P. R. (1970). State-trait anxiety inventory (STAI). *BiB 2010*, page 180.

Spink, J. (1993). *Developmental riding therapy.* Therapy Skill Builders.

Staudte, M. and Crocker, M. (2008). The utility of gaze in spoken human-robot interaction. In *Proceedings of Workshop on Metrics for Human-Robot Interaction 2008, March 12th*, pages 53–59. Citeseer.

Stiehl, W. D., Lieberman, J., Breazeal, C., Basel, L., Lalla, L., and Wolf, M. (2005). Design of a therapeutic robotic companion for relational, affective touch. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop*, pages 408–415. IEEE.

Suga, K., Sato, M., Yonezawa, H., Naga, S., Shimizu, J., and Morita, C. (2002). Change in the concentration of salivary IgA by contact of elderly subjects with a pet robot. *International Journal of Analytical Bio-Science*, 25:251–254.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11(9):1004–1006.

Sung, J.-Y., Guo, L., Grinter, R. E., and Christensen, H. I. (2007). *My Roomba is Rambo: intimate home appliances*. Springer.

Syrdal, D. S., Dautenhahn, K., Woods, S., Walters, M. L., and Koay, K. L. (2006). 'doing the right thing wrong' – Personality and tolerance to uncomfortable robot approaches. In *Robot and Human Interactive Communication. ROMAN 2006. 15th IEEE International Symposium*, pages 183–188. IEEE.

Tamura, T., Yonemitsu, S., Itoh, A., Oikawa, D., Kawakami, A., Higashi, Y., Fujimooto, T., and Nakajima, K. (2004). Is an entertainment robot useful in the care of elderly people with severe dementia? *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 59(1):M83–M85.

The European Commission (2012). *Special Eurobarometer 382: Public attitudes towards robots.* Retrieved August 1, 2016 from http://ec.europa.eu/publicopinion/archives/ebs.

Thill, S., Pop, C. A., Belpaeme, T., Ziemke, T., and Vanderborght, B. (2012). Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn, Journal of Behavioral Robotics*, 3(4):209–217.

Toufexis, A., Gorman, C., and Wormley, M. (1987). Furry and feathery therapists. *Time*, 129(13):74–74.

Uchino, B. N., Cacioppo, J. T., and Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin*, 119(3):488.

UN (2012). *World population prospects: The 2012 revision.* United Nations, New York.

US Surgeon General (2001). Culture counts: The influence of culture and society on mental health. *US and Center for Mental Health Services.*

Valdez, P. and Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4):394.

Van Widenfelt, B. M., Treffers, P. D., De Beurs, E., Siebelink, B. M., and Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*, 8(2):135–147.

VProBacklight (2016). *The human robot.* Retrieved February 2, 2016 from https://www.youtube.com/watch?v=sXz0boNmwak.

Wada, K. and Shibata, T. (2006). Living with seal robots in a care house-evaluations of social and physiological influences. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference*, pages 4940–4945. IEEE.

Wada, K. and Shibata, T. (2007). Living with seal robots: Its socio-psychological and physiological influences on the elderly at a care house. *Robotics, IEEE Transactions on*, 23(5):972–980.

Wada, K., Shibata, T., Musha, T., and Kimura, S. (2005a). Effects of robot therapy for demented patients evaluated by eeg. In *Intelligent Robots and Systems: (IROS 2005) IEEE/RSJ International Conference*, pages 1552–1557. IEEE.

Wada, K., Shibata, T., Sakamoto, K., and Tanie, K. (2005b). Quantitative analysis of utterance of elderly people in long-term robot assisted activity. In *Robot and Human Interactive Communication. ROMAN 2005. IEEE International Workshop*, pages 267–272. IEEE.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896.

Walsh, F. (2009). Human-animal bonds i: The relational significance of companion animals. *Family Process*, 48(4):462–480.

Walters, M. L., Dautenhahn, K., Te Boekhorst, R., Koay, K. L., Syrdal, D. S., and Nehaniv, C. L. (2009). An empirical framework for human-robot proxemics. *New Frontiers in Human-Robot Interaction.*

Wang, Y. and Ollendick, T. H. (2001). A cross-cultural and developmental analysis of self-esteem in chinese and western children. *Clinical Child and Family Psychology Review*, 4(3):253–271.

Wei, M., Russell, D. W., Mallinckrodt, B., and Vogel, D. L. (2007). The Experiences in Close Relationship Scale (ECR) - short form: Reliability, validity, and factor structure. *Journal of Personality Assessment*, 88(2):187–204.

Weller, J. A., Shackleford, C., Dieckmann, N., and Slovic, P. (2013). Possession attachment predicts cell phone use while driving. *Health Psychology*, 32(4):379.

Wemelsfelder, F., Hunter, A., Paul, E., and Lawrence, A. (2012). Assessing pig body language: Agreement and consistency between pig farmers, veterinarians, and animal activists. *Journal of Animal Science*, 90(10):3652–3665.

Wikipedia (2016). *AIBO.* Retrieved February 2, 2016 from https://en.wikipedia.org/wiki/AIBO.

Wilson, E. O. (1984). *Biophilia.* Harvard University Press.

Winnicott, D. W. (1953). Transitional objects and transitional phenomena. *The International Journal of Psycho-Analysis*, 34:89.

Wolff, A. I. and Frishman, W. H. (2004). Animal-assisted therapy in cardiovascular disease. In *Seminars in Integrative Medicine*, volume 2, 4, pages 131–134. Elsevier.

Yilmazyildiz, S., Henderickx, D., Vanderborght, B., Verhelst, W., Soetens, E., and Lefeber, D. (2013). Multi-modal emotion expression for affective human-robot interaction. In *Proceedings of the Workshop on Affective Social Speech Signals (WASSS 2013), Grenoble, France.*

Zilcha-Mano, S., Mikulincer, M., and Shaver, P. R. (2011). Pet in the therapy room: An attachment perspective on animal-assisted therapy. *Attachment & Human Development*, 13(6):541–561.

Zotterman, Y. (1939). Touch, pain and tickling: An electro-physiological investigation on cutaneous sensory nerves. *The Journal of Physiology*, 95(1):1.

Zunt, D. (2002). *Who did actually invent the word robot and what does it mean?* Retrieved January 3, 2016 from http://capek.misto.cz/english/robot.html.