An exploration of factors influencing human responses to short term interaction with a pet robot

Raihah Aminuddin

The University of Sheffield
Faculty of Engineering
Department of Computer Science

A thesis submitted for the degree of Doctor of Philosophy
2019

Supervisor: Dr. Amanda Sharkey

“Take care of your souls.” - Al Ma’idah 5:105 -

Contents

List of Figures

List of Tables

Abstract

Past research has found that intervention with Paro, a seal robot, on healthy adults can improve psychophysiological responses (Mitsui et al. 2001*a*,*b*, Kawaguchi et al. 2012). According to Bethel & Murphy (2010), psychophysiological study focuses on understanding human emotion and behaviour during an event through observation, physiological and psychological measures. However, these previous studies of Paro were limited by small participant sizes (less than 10 participants), and did not investigate the effectiveness of Paro in reducing stress such as environmental stress. Furthermore, little is known about the circumstances in which interacting with Paro can help to reduce stress as indicated by changes in psychophysiological responses.

Three studies were conducted in laboratory settings, in order to limit any extraneous variables that can affect the psychophysiological responses. This thesis aims to identify factors that can influence psychophysiological responses to short-term interaction with a Paro in the healthy adult population. It also aims to understand how the factors affect the effectiveness of Paro intervention in reducing effects of stress. The psychophysiological responses were measured using self-reports, physiological sensors (skin conductance responses, heart rate, and heart rate variability) and video recording. The stress used in this thesis is induced environmental noise.

A study was conducted (n = 76) to explore which features of Paro (active and inactive) and of human behaviour towards the robot (talking and stroking) were responsible for any effects it creates. It was found that interacting with Paro aroused skin conductance responses and its presence during the intervention increased positive moods. At the same time, the positive moods influenced the effectiveness of the Paro intervention in reducing stress.

A subsequent study (n = 104) compared the effects of interacting with an active Paro with the effects of stroking a furry bolster, and explored the stress reducing effect of Paro intervention as a pre-stress treatment. The novelty effect of Paro in the intervention was also investigated. It was found that the Paro intervention was not enough to minimize and reduce the impact of induced stress. However, there was a trend that suggests the Paro intervention

has more impact in lessening the stress than the bolster intervention as indicated by improvements in physiological responses (such as heart rate and heart rate variability). At the same time, it was found that novelty affected how participants interacted with the Paro (based on video observation during the intervention), but did not change the positive effect of the Paro intervention.

A final study (n = 104) explored two intervention formats: either Paro or bolster as pre- and post-stress intervention. This study also explored the effectiveness of Paro or bolster intervention using a mathematical task. It was found that the Paro intervention was more suitable as a post-stress intervention, as suggested by improvement in physiological responses (heart rate and heart rate variability). However, no improvement in math performance was found in either Paro or bolster interventions.

Additionally, in this study and Study 2, the bolster intervention showed the importance of stroking soft fur in reducing stress and negative moods.

There are five key findings of this thesis: Large samples of data were obtained from (i) healthy young adults (n = 284). Paro intervention is more effective as (ii) post-stress treatment, in terms of reducing effects of (iii) induced environmental stress in (iv) healthy adults. The thesis also found (v) factors that influenced the changes in psychophysiological responses and the effectiveness of Paro interventions such as stressors and intervention formats.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Acknowledgements

I would like to thank my supervisor Dr. Amanda Sharkey for all her probing questions and insightful comments, patience and enthusiasm throughout the years of my PhD.

I would like to thank Dr. Liat Levita and Nat from the Psychology department for their help in setting up the lab and the equipment for the data collection. I would also like to acknowledge all of my volunteers and friends. Thank you for your commitment and time for the research.

I am also very grateful to my family for all their love and support. Especially to my lovely sisters, for endless text and video chats, for their willingness to listen to my problems and complaints, this journey would have been very lonely without them. I would like to particularly thank my friend Bell, without her I would still be lost in Sheffield. I would like to thank my friends for keeping me from being insane and for always being there when I am in crisis or need: Zirah, Ena, Yenyen, Aina Putri, Nor, Kirah and family, Rabab, Erfan, Fra and Aditi.

Although, this journey has been the most incredible and exciting experience in my life, it also has been extremely challenging and filled with self-doubts, self-criticism and fear of rejection. A special mention for my adopted pet, Manis, for being a good companion during my thesis writing. Thank you for always reminding me to take a few minutes break and to play with her.

I would like to dedicate this thesis to my strong and lovely aunt, who died in battle with cancer on 2nd March 2018.

Publications arising from this research

Aminuddin, R., Sharkey, A. & Levita, L. (2016) Interaction with the Paro robot may reduce psychophysiological stress responses. In ACM/IEEE International Conference on Human-Robot Interaction, New Zealand, April 2016, pp. 593 - 594.

Aminuddin, R. & Sharkey, A. (2016) Does interacting with a robot reduce your physiological stress response? In New Friends 2016 Conference, Barcelona, Spain, pp. 4 - 5.

Aminuddin, R., & Sharkey, A. (2017) A Paro robot reduces the stressful effects of environmental noise. In Proceedings of the 35th European Conference on Cognitive Ergonomics 2017 (ECCE 2017). Umeå, Sweden, September 19 - 22, 2017. ACM, New York, NY, USA, pp. 63-64.

Chapter 1

Introduction

The number of people of all ages that are suffering from mental health problems such as anxiety, stress, depression has been increasing yearly especially among the younger generation (i.e. the student population) (World Health Organization 2010, Kumaraswamy 2013, World Health Organization 2014). If stress is not being treated properly, it can lead to high levels of stress, anxiety, loneliness, and nervousness (Robotham 2008, Perou et al. 2013). It is very important for people who suffer from mild to severe mental health problems to seek advice and therapy to reduce or buffer the stress effect (American Psychological Association 2015). According to Macaskill (2013), there is an increasing number of people looking at non-pharmacological stress reducing treatment. One of the treatments is pet therapy Beetz et al. (2012a).

Stress reducing intervention and animal therapy:

There is an empirical evidence of the effectiveness of animal therapy in reducing psychophysiological of stress (Fine 2010, Lundqvist et al. 2017). However, there is also a considerable debate about its effects and about the use of animal therapy, for example, ethical concern about animal welfare, cost, and potential harm to humans (Wilson & Barker 2003, Melson et al. 2005). The emergence of advanced technology such as robotic pets provides an alternative to animal therapy.

Pet robots and Paro robot:

Past research showed that a robotic pet could provide similar benefits to those of animal therapy by reducing stress responses (Bemelmans et al. 2012, Chang & Sung 2013, Kachouie et al. 2014, Preuß & Legal 2017, Broadbent 2017). One of pet robots that has been used commercially is Paro, a seal robot. Based on literature, there were evidences of short- and long-term intervention with Paro showed an improvement in social behaviour in elderly people with Dementia and autistic children (Rabbitt et al. 2015, Abdi et al. 2018, Moerman et al. 2018). Many studies claimed that Paro acts as mediator but not as a tool to replace human in social interaction. The presence of Paro during the intervention not only trigger communication between elderly people in nursing home but helps the elderly to improve their cognitive memory. Improvement in cognitive memory help the elderly to recollect their past events or old memory. This also help to improve their short-term memory and enhanced their communication skills, hence reduced loneliness in elderly people. The intervention with Paro also helps to reduce pain and anxiety in hospitalized patients with diagnosed diseases (Okita 2013). There are also studies that mentioned that the intervention with Paro helps in reducing stress in terms of psychophysiological measures such as improving in moods and physiological functions (Mitsui et al. 2001*a*,*b*, Carretero & Sáez 2017, Crossman et al. 2018).

Gap in the literature of Paro:

Indeed, there is empirical evidence that interacting or intervention with Paro may help to reduce stress, pain and minimizing the symptoms of Dementia. At the same time, the intervention helps to improve quality of life of people with Dementia. Therefore, the potential of Paro interventions in reducing stress and improving moods should not be limited to those population. The interventions should also be applied and benefits by population such as healthy young adults i.e. students. However, it still unclear whether the Paro interventions might help to reduce stress in healthy young adults, especially in improving mental and physical health.

There is a large gap in literature review of Paro when investigating the effectiveness of Paro intervention among healthy young adults and involving methodological issue. The issues such as a small sample size of participants, lack of controlled condition and within-design study and no study understanding the effects of Paro as pre- or post-stress treatment, especially in terms of short-term stress reducing intervention.

At the same time, there was no research of Paro investigating the effects of Paro intervention in reducing effect of stress that related to environmental stress i.e. noise. There was some evidence that environmental stress is one of the factors that increase the risk of cardiovascular diseases (Walker et al. 2016), increase the symptoms of agitation in people with Dementia (Dewing 2009) and worsen performance when doing a task (Szalma & Hancock 2011). Therefore, it is important to seek an effective intervention that can help to minimize or reduce the effect of stress causes by exposure to noise.

1.1 Relevance of this thesis

The thesis begins with understanding and filling a gap in Paro literature, based on Mitsui et al. (2001*a*) study. The relevance of this study is that it provides strong evidence of the effect that a Paro intervention has on improving psychological and physiological functions with a large number of participants (based on healthy adults population) and a rigorous study design.

In terms of the stress reducing intervention, it is important not only to understand the effectiveness and final outcome of the interventions, but also to understand the immediate effects during the interventions. At the same time, it is important to explore factors such as robot features or human behaviour towards a robot. These factors can help to understand the impact of the intervention on reducing stress and help to improve the design of the intervention.

This study not only relevant to study of stress reducing intervention, at the same time, it is also relevant to understanding the effect of acute stress on psychological and physiological responses. It was acknowledged that the changes in psychological and physiological responses during stress can provide a clue how an individual responds to and copes with the stress (Walker et al. 2016, Fish et al. 2018). However, it still is unclear how the stress reducing intervention of Paro creates any changes on psychological and physiological responses especially in short-term intervention settings. Furthermore, the use of physiological and psychological measures together can contribute on understanding the human experience that cannot be reported using self-reports.

1.2 General aims

The general aims of this thesis are as follows:

1. To explore factors in short-term interaction with Paro that affect human psychological and physiological responses. Those factors include:
   - the ways of interacting with the Paro robot.
   - the Paro robot features themselves, and at the same time, to compare the Paro features with inactive object such as bolster.
   - different intervention formats: (i) pre- and (ii) post-stress reducing intervention.
2. To compare the effects of a Paro session and a stress session on psychological and physiological responses.
3. Finally, to investigate whether the interaction with a Paro reduced the effects of different types of induced stress.

1.3 Main research questions

The following are the main research questions for this thesis:

1. What are the immediate effects of short-term interaction with a Paro on observed facial expression, self-reported and physiological responses?

2. Does an intervention session with Paro have an impact on reducing the effects of different forms of induced stress on self-reported and physiological responses?

3. What effects does the implementation of different intervention formats (i.e. either interacting with Paro before or after a stress inducement session) have on self-reported and physiological stress responses?

4. Does interacting with Paro have a different impact on psychological and physiological responses compared to stroking a bolster?

1.4 Three studies in this thesis

Based on the general aims outlined in the previous section, a summary of the purpose of the three studies reported in this thesis follows:

Study 1:

Study 1 aims to understand the effects of interacting with Paro and its effect on induced stress. Additionally, Study 1 investigates aspects of human behaviours towards the robot such as talking and stroking, in order to know which behaviour induces more positive effects during interaction with Paro and which reduce the effects of induced stress. This study also explores the effects of robot behaviour such as being active or inactive and their contribution towards creating positive moods during the interaction and reducing induced stress. In Study 1, stress was induced by means of exposure to an environmental noise. Participants were exposed to noise, followed by a session with Paro.

The responses were collected using self-reported positive and negative affect, stress and physiological measurement: skin conductance responses, and using observed facial expression.

Study 2:

Based on the results of Study 1, a further investigation into the effects of interacting with

Paro and its effect on induced stress was undertaken. In this study, a comparison between interacting with an active robot (Paro) versus an inactive object (a bolster) was made, in order to determine which aspects of Paro are responsible for the effects that it creates. In this study, the intervention format was changed, in order to investigate the effects of interacting with Paro without the effect of a previous stress induction. A session with the Paro or a bolster was followed by a stress inducing noise session. This made it possible to explore the potential of interacting with Paro on minimizing or reducing the effects of subsequent induced stress. This study also included an investigation of the effects of familiarity with Paro prior to the intervention session.

The responses were collected using self-reported positive and negative affect, stress and emotions, and physiological measurement: skin conductance responses, heart rate and heart rate variability, and using observed facial expression.

Study 3:

In Study 3, a comparison between the effects of interacting with Paro and a session with a bolster was further investigated. However, Study 3 also explores the stress reducing effects of intervention in different formats. There are two formats of intervention in Study 3: an intervention session followed by a stress session, and a stress session followed by an intervention session. At the same time, based on the different intervention formats, Study 3 examined whether an earlier intervention session with Paro or a bolster helps to reduce stress. At the same time, the study examined whether a Paro or a bolster session may improves mathematical performance compared to no earlier intervention session with Paro or bolster. Furthermore, Study 3 also compares the difference in the effects of a stress session and an intervention session, in order to reveal the different psychological and physiological responses under two different circumstances.

The responses were collected using self-reported positive and negative affect, stress, emotions, perceived stress and mathematical task scores, and using observed facial expression. Then the physiological responses were measured using skin conductance responses, heart rate and heart rate variability.

1.5 Thesis contributions

By addressing the research questions listed in the previous section, this thesis is expected to enrich the knowledge of the short-term effects of interacting with a pet-robot on psychological and physiological responses. The contributions of this thesis are as follows:

Contribution 1: This is the first study to look at the short-term effects of interacting with a Paro robot on induced environmental stress (in Study 1 and Study 2), and a combination of environmental stress and arithmetic stress test (in Study 3). This helps in understanding the potential effects of Paro on reducing different daily stressors.

Contribution 2: The thesis extends previous research with older adults with cognitive impairment to examine the effects of Paro on a non-medical population of healthy adults. All the studies in this thesis were conducted in a laboratory setting using a rigorous experimental methodology that looks at both self-report (psychological) and physiological responses.

Contribution 3: This thesis explores the factors that influence the effects of intervention with Paro as a positive intervention, as measured by the psychological and physiological responses. The factors include different ways of interacting with Paro (talking versus physical interaction), familiarity with Paro before a therapy session, and a comparison of the effects of physical interaction with an active (Paro robot) and an inactive object (bolster). Then based on the immediate effects during intervention, this thesis explored its potential effects in reducing the effects of different types of induced stress.

Contribution 4: At the same time, this thesis extends an understanding of how a Paro intervention and a stress session influence changes in skin conductance response (SCR), heart rate (HR) and heart rate variability (HRV), in terms of interaction and stress recovery. An understanding of the changes of physiological responses is important, in order to assess the effectiveness of therapy as mechanism of stress prevention or recovery especially in a clinical situation.

Contribution 5: This thesis extends previous research of Paro in healthy adults by exploring the effects of different intervention formats based on a manipulation of the timing of the session. Comparing the intervention formats (stress inducement before or after a session with Paro), may indicate the potential use of pet robots for reducing the effects of anticipated upcoming stressors or for reducing the effects of being exposed to a stressor. For example, in a real case situation such as in a waiting room before meeting a doctor, or to comfort people in temporary shelters, who were affected by tsunami in Japan.

1.6 Thesis outline

This thesis consists of six chapters.

Chapter 2 reviews previous studies of animal therapy and their advantages and disadvantages. Then, it continues to introduce studies of pet-robot therapy that have been used as mediator in social interaction, companion and provide assistance to user. The chapter also discusses previous studies of Paro. There is evidence that interacting with Paro can improve social interaction, reduce pain (by decreasing pain medication intake), decrease anxiety and loneliness, and also enhance psychophysiological responses. The review includes the study of Paro on a variety of populations (i.e. children, young adult, and elderly people). Then, a summary of the effects of Paro is given, followed by a discussion of the limitations of previous studies and the rationale for this thesis.

The first study in Chapter 3 focuses on exploring the aspects of Paro and human behaviour towards the robot that may help to alleviate the effects of induced stress. This study also explores the immediate effects of Paro intervention can have on improving moods and physiological stress responses.

Based on the result of the first study, the second study in Chapter 4 examines the stress-reducing effects of interacting with Paro on subsequent stress induction. The intervention with Paro also compare with the intervention with a bolster and resting alone. The intervention of Paro also been compared with a factor such as familiarisation with Paro in an introduction session. This is an important consideration as novelty effect in human-robot interaction study could influence moods. At the same, the factor could also influence the effectiveness of Paro intervention in minimizing the impact of stress.

The results of the first and second studies inform the design of the final study in Chapter 5. The study investigates the effects of different intervention formats on influencing how the intervention helps in reducing participants' stress as a responses to induce stress. At the same time, the study explores more details on comparison between different format interventions (Paro intervention as pre- and post-stress treatment) In addition, the study includes a mathematical task during the stress session, in order to look at the effectiveness or Paro intervention. The study also compare the effects of interacting with Paro and stroking a bolster.

Finally, Chapter 6 discusses the answers to the main research questions of this thesis along with implication of this thesis on Paro literature. The key achievements of this thesis: (i) a large sample of data collected from healthy adults population. (ii) Paro intervention is more effective as a post-stress treatment to reduce effects of environmental stress. (iii) Provide more understanding on the factors that influenced the changes of psychological and physiological during stress and intervention session. The chapter also reviews the limitations of this study and highlights some suggestions for future studies. In future studies, other factors that could influence the effectiveness of intervention such as human behaviour i.e. hugging and talking (more specifically looking at the effects of reminiscence) should be done.

Chapter 2

Literature Review

The chapter begins with a general review of pet or animal therapy, or more specifically, the benefits and limitation of animal therapy. Then, in section 2.3, the effects of Paro robot therapy are reviewed. Next, in section 2.4 and 2.6, a review of stress and methods of assessing the emotions in human-robot interaction is provided. In further sections highlight a review on methods of analysis for physiological measures (section 2.6.4) and how to perform data analysis in statistics (section 2.7). Then, the chapter concludes with the rationale of this thesis based on the reviews on Paro and stress literature (in section 2.8).

2.1 Animal therapy

The term "pet therapy" was formally used in 1960s (Hooker et al. 2002). Animal therapy provides methods to heal the human mind and body. Animal therapy can decrease loneliness and boredom (Levinson 1962).

A study by Friedmann et al. (1980) found that heart disease patients who owned a pet had a higher survival rate compared to non-pet owners. The study was also supported by Friedmann & Thomas (1995), who found that patients with myocardial infarction (i.e. cardiovascular disease) who owned a pet lived a year longer than non-owners. Barker et al. (2010) also found a relationship with pet-ownership in reducing physiological stress responses and self-reported stress compared to when interacting with unfamiliar pet. The ownership also related to attachment behaviour between the pet and owner which contribute to positive and stress reducing effects during therapy session.

An additional study by Nagasawa et al. (2009) also investigated attachment behaviour. They found that a gaze exchange between owner and pet for a few minutes increased oxytocin. The oxytocin is a hormone produced in a small part of the brain, the hypothalamus. The oxytocin hormone releases helped to stimulate the interaction between pet and owner. Furthermore, pet therapy also showed positive effects in preventing cardiovascular disease (Anderson et al. 1992).

A study by Shiloh et al. (2003) found that petting a pet has an effect in reducing induced anxiety. Braun et al. (2009) also found a reduction of pain levels in hospitalised children after a short session with a dog. A study by Tsai et al. (2010) found a decrease in heart rate after children engaged with animal therapy during hospitalization. This is because hospitalized children will have a higher heart rate cause by pain and anxiety.

However, in contradiction to the previous studies described above, a study by Palestrini et al. (2017) did not find changes in heart rate or behaviour across sessions in a paediatric hospital. Patients showed increased breathing rates, but, the quick breathing was linked to the high temperature set in the ward.

Lass-Hennemann et al. (2014) used a service dog as a companion to reduce stress rating responses during acute stress. The acute stress was induced by traumatic video which contained a traumatic rape scene. The study was conducted in a laboratory with a session length of 20 minutes. The participants were female students (n = 80) and the study used a randomised control study design. There were four conditions in the study: participants were accompanied with (i) dog, (ii) stuffed dog, or (iii) a friendly person, or (iv) nothing while watching the stressful video. However, the study did not find any changes in physiological responses i.e. cortisol, heart rate and blood pressure across all the conditions. The study also did not find an increase in self-reported positive affect across conditions.

At the same time, animal therapy has been found to increase social interaction between participants in elderly care homes and in hospitals (Broadbent 2017). The participants were nurses, staff, patients, and visitors. Animal therapy also showed improvement on human psychophysiology. For example, Calvo et al. (2016) found that animal therapy resulted in improved negative self-reported responses and cortisol levels after a few sessions on patients with schizophrenia. Additionally, several studies looked at the effects of watching aquatic plants or animals in an aquarium in reducing stress (DeSchriver & Riddick 1990, Edwards & Beck 2002, Barker et al. 2003). However, none of these studies found a significant effect when watching the aquarium on heart rate and blood pressure. This could due to a lack of physical interaction (Beetz et al. 2012a). Contrary to the results of a study by Edwards & Beck (2002), the authors found an increase in healthy diet intake in patients with Alzheimer's disease when the aquarium was placed in a dementia unit.

Despite the positive effects of animal therapy on humans, there are several ethical concerns about animal therapy such as (i) animal harm, (ii) cost and (iii) diseases or harm of animals.

Animal Harm:

There are some concerns about the use of animals in therapy programs, such as the way of interacting with the animal, when it was reported that the animal had been treated badly by the patients and staff in the unit (Heimlich 2001, Ehrén 2014). Ethical concerns have also been raised about the use of pet therapy (i.e. dogs, cats and horses), where the animal showed signs of fatigue due to the environment in the therapy room (i.e. high room temperature, more than an hour per session) (Iannuzzi & Rowan 1991). The animal may have shown signs of stress due to the interaction with unfamiliar or different people. Due to interaction with unfamiliar people, an animal can become aggressive (Mongillo et al. 2015). However, there are some guidelines to ensure the well-being of the animal (Wohlfarth & Sandstedt 2016). The staff should be properly trained on how to handle any situation (Serpell et al. 2010).

Another concern was anxiety in animals such as anxiety in dogs. The anxiety in dogs can be due to separation from other dogs. Dogs in therapy programs should be able to behave naturally and mix with other dogs (Schwartz 2003, Zamir 2006). An attachment to a pet can increase the positive social interaction between owner and pet, however, it can also result in increased anxiety and distress when the animal is separated from the pet owner (Flannigan & Dodman 2001, Schwartz 2003). A similar effect can occur in the pet owner, and the loss of a pet can cause grief (Beck & Katcher 2003). Reduced heart rates were also found in dogs with separation anxiety (Wormald et al. 2017).

Cost:
The cost of taking care of pets can be both expensive and time-consuming. The care cost of an animal pet can range from 8000 to 10,000 dollar a year (Herzog 2011). The animals need proper care that includes food, home and medical treatment. Looking after an animal could be a problem for elderly people with limited physical movements or cognitive ability (Beck & Katcher 2003).

Diseases or harms of animals:
An animal can transfer chronic or contagious diseases such as scabies, fleas and ringworm (Hooker et al. 2002, Brodie et al. 2002, Lutwack-Bloom et al. 2005). Elderly people or patients or staff in medical settings could be potentially exposed to these harmful diseases (Barba 1995, Sorrell et al. 2010). The animal also can bite or scratch the owner (Brickel 1979). The fur of an animal can cause or worsen the symptoms of asthma and allergies (Brodie et al. 2002, Yawn 2008).

Overall, without doubt, a large number of studies of animal therapy have shown an improvement to human health and social interaction. However, due to some limitations of animal-assisted therapy, the use of pet-robot therapies are gaining popularity and have received positive reactions from elderly people and children (Broadbent et al. 2009, Bharatharaj et al. 2017). Many studies of pet-robot therapy in human-robot interaction that aim to understand human behaviour or emotions towards robot follow a similar study design as those used in animal-assisted therapy. The next section will review pet-robot therapy and then we will focus on the reviews of previous studies that related to Paro robot.

2.2 Pet-robot therapy

Pet-robot therapy has been shown to provide similar benefits as animal therapy. Most of intervention studies of pet-robot therapy have been adapted from animal therapy. For example, interacting with a pet-robot has been shown to lower blood pressure (Robinson et al. 2013c) and reduce anxiety (Okita 2013). Shibata (2012) also stated that the use of pet-robot therapy in care homes can increase social activity among patients. Robot therapy also can be used as a means to distract a patient from painful experience (Okita 2013). Pet robot therapy not only can have a positive impact on patient and staff in hospital, but can also affect visitors emotionally. As a result of using an pet-robot as a mediator, visitors felt relaxed and found the visit more enjoyable (Gelderblom et al. 2010). In addition, the use of pet-robot therapy can reduce the time needed by staff to take care of a real animal and reduce possible threats to human posed by real animals such as allergies and being bitten.

There are a number of studies of pet-robot therapy that have been investigated, following are some of the examples of the studies where the pet-robot act as a human companion.

A long-term interaction study by Melson et al. (2005) found that AIBO[1], a dog robot, can have a similar effect of companionship as a real dog. The study had two conditions: (i) AIBO and (ii) trained dog. The study was conducted in preschool settings. 80 participants were divided into three ages groups: 7 to 9, 10 to 12, and 13 to 15 years. The children had approximately 20 minutes session with AIBO then another approximately 20 minutes session with a real dog accompanied by a researcher. However, the older children group (13 to 15 years) had an alone 5 minutes session with AIBO, and another 5 minutes session with a real dog. Then, they were interviewed by the researcher for 35 minutes. Based on the results

[1] The official website of AIBO robot: https://us.aibo.com/

of observation and interviews, the children prefer to play with the real dog than with AIBO. However, the children treated the AIBO similarly to the real dog.

A study by Gustafsson et al. (2015) looked at the benefits of JustoCat[2], a robotic cat and the study found reduced agitation and improved quality of life of elderly people (n = 4). The JustoCat was developed based on reminiscence therapy framework. The JustoCat is an interactive robot that considered an importance of hygiene while using robot therapy with fur in a care home. The fur is washable, in order to keep elderly people from the risk of infection. An intervention study was conducted for 10 weeks by recruiting four elderly people with severe dementia in a Swedish care home. In the experiment, there were two stages, one was intervention for elderly people and other was an interview with relatives and caregivers. The interview was conducted to assess opinion on the use of robotic cat in care home. The results showed that the robot resulted in an improvement in the elderly people's behaviour.

Another study by Moyle et al. (2016) focused on the use of CuDDler, a teddy bear robot for elderly people with dementia. They looked at its effects on agitation behaviour and social improvement. Case studies of five elderly people with dementia in a care home were collected. The CuDDler was programmed to perform a response to participant physical interaction such as patting, and stroking. In this study, participants could freely interact with the robot for an hour per week for 5 weeks. Their behaviour was video-recorded. The data form video was analysed using Observed Emotion Rating Scale. Semi-structured interviews were also conducted at the end of the experiment. Based on data collection from the interview, the results showed that it is important to consider about the size and weight of the robot when developing it especially when the end users are elderly people. However, the CuDDler showed a potential to act as a mediator in social interaction. The participants' interactions were shown to have contributed to improvements in their mood.

There was a recent on-line article by Cooper (2018) about a cuddle pillow robot named Somnox[3]. The robot acts as a sleeping companion where people can cuddle it, and the robot

[2]The official website of JustoCat robot: http://www.justocat.com/
[3]The official website can be found on https://meetsomnox.com/

provides comfort through a sound of heart beats or sounds of nature, or lullabies. The sound can be chosen based on personal preference. At the same time, the pillow robot was also able to monitor the sleeping patterns of the person cuddling it. It was claimed that the pillow robot was able to reduce the symptoms of stress, and hence improving the quality of sleep (Somnox 2017). But, a review by Boxall (2018) in a recent on-line article, the author mentioned that he found that the pillow robot is less likely to replace a human or a dog as sleeping companion. But, it might be not in the case of a night spent with the robot. In a white paper published by the company in 2017, Somnox (2017) found that most participants showed an improvement in average time of sleeping. In the study, participants ages between 25 to 55 years (n =10) were asked to sleep accompanied by the pillow robot for 5 nights in their own home. Then on the last day of experiment, participants were asked to fill in an on-line questionnaire that asking about their sleeping pattern. This findings showed that there are some potential that the pillow robot can be used for healthy adults and elderly people.

Among these pet-robots, Paro has been the focus of a significant number of studies investigating its therapeutic effects on social interaction, health and mood improvement. In the following section, the effects of Paro will be reviewed. Paro's effects are categorised as: (i) Improved social interaction, (ii) reduced pain, anxiety and loneliness, and (iii) psychophysiological improvement. Reviews of Paro will be discussed in the next section.

2.3 The therapeutic effects of Paro the robot seal

Paro, a seal therapeutic robot can be categorized as a companion-type robot and is illustrated in Figure 2.1. Paro has been commercially available in Japan since 2005. It has been marketed to more than 30 countries including Europe, Korea, Denmark, the Netherlands and the United States. The robot has been used as a therapy tool for patients in hospitals and care homes. Paro also has Food and Drug administration (FDA) approval as a medical device in the United States. Paro has been used in clinical studies in more than 10 countries (Yang 2015). It has been claimed that Paro can stimulate the brain activities of elderly people with dementia when interacting with it (Shibata 2012).

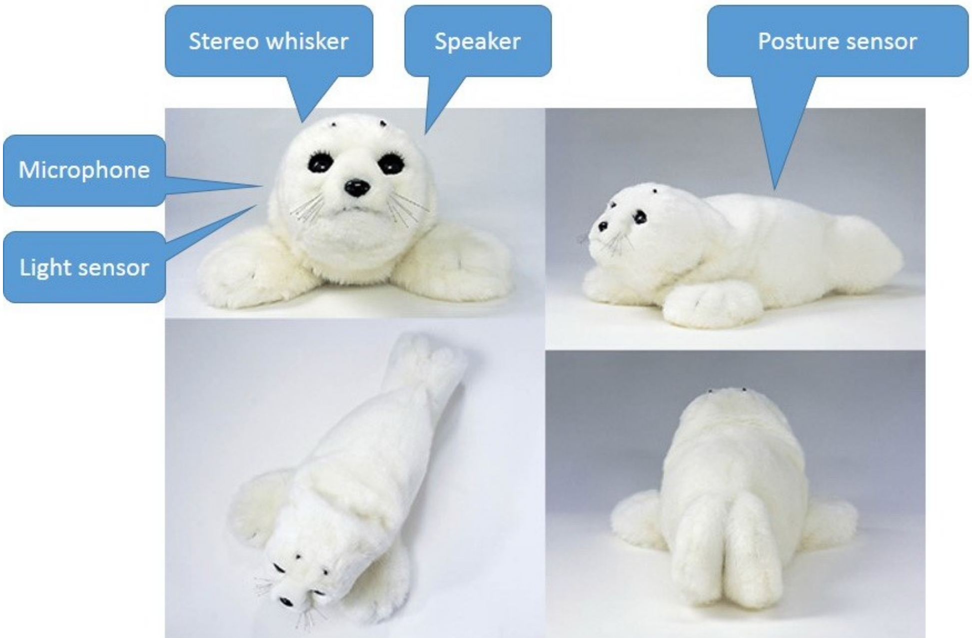


Figure 2.1: Paro robot, from (Center 2015)

Paro has a physical shape similar to that of a baby harp seal. Paro's body is covered with soft white fur. The tactile sensors located between the fur and endoskeleton are used to measure human contact with Paro. Paro shows sad and cute expressions with movements of the eyelids. There are four sensors used in Paro including (i) light sensor, (ii) microphone (for speech recognition and hearing), (iii) body area with tactile sensors and (iv) temperature. The (i) light sensors are located on Paro's eyes and are able to detect any movements. For example, when a human moves in a different direction, the Paro will move its head towards the human. Another sensor is (ii) microphone for hearing and speech recognition. The speech is captured by a microphone located near Paro's nose. It is known that Paro can recognize its name when called and understand some English and Japanese words. Paro reacts to unexpected events such as a loud noise and is able to locate the source of the noise. The (iii) tactile sensors can detect touch. Paro has a similar (iv) body temperature to a real animal (Shibata et al. 2011, Wada et al. 2013). The body temperature produce by internal heating system inside Paro's body.

Other important functions that Paro has are two types of behaviour: (i) proactive and (ii) reactive. The proactive behaviour has two layers: (i) behaviour-planning layer and (ii) behaviour-generation layer. Both layer contributes to behaviour of Paro, the behaviour system of Paro focusing on making the Paro reacts like the real seal. At the same time, create unpredictable movements and sounds which is asking and begging for an attention.

The reactive behaviour of Paro is the ability to make a little movement with his neck and rear paddle. For example, if the Paro hear a sound, it will move its head towards the source of the sound.

Paro has reinforcement learning for long-term memory. Its reinforcement learning is said to give it the ability to react to the positive and negative values of interactions such as stroking or undesirable hitting. The detection used the tactile sensors located underneath the soft fur of Paro. Reinforcement learning is claimed to help Paro's behaviour change slowly according to the behaviour of its owners (Shibata & Coughlin 2014). Paro also has the ability to remember the new name given by its owner (Wada et al. 2013). However, owners cannot change the Paro's behaviour manually. The physiological behaviour of Paro is similar to a real seal animal in the sense that it has diurnal rhythms for morning, daytime and night. Paro sleeps at night and when it is tired or needs to be recharged. The Paro is recharged using a battery charger that looks like a pacifier (Wada et al. 2005a, 2008).

Table 2.1 summarizes the characteristics of Paro.

Table 2.1: Paro's characteristics

Characteristics	
Body	Soft fur, eyelid,endoskeleton
Sensors	(i) Light, (ii) Speech recognition and hearing, (iii) Touch sensors and body temperature
Proactive behaviour includes behaviour-planning and -generation layer.	Internal states of Paro and making the Paro move like a real animal such as the movement of the actuator.

Continued on next page

Table 2.1 – continued from previous page

Characteristics	
Reactive behaviour	Give reaction to spontaneous interaction with owner. For example, the ability to recognize their name when being called.
Physiology	Diurnal rhythm behaviour, sleeping during night.

A number of studies have found that Paro provides positive effects on humans such as (i) improved social interaction (Wada et al. 2004*b*, Marti et al. 2005, Taggart et al. 2005, Kidd et al. 2006, Wada & Shibata 2007b, Yamamoto & Kimura 2007, Pipitpukdee et al. 2011, de Graaf & Allouch 2014, Takayanagi et al. 2014, Sung et al. 2014, Yu et al. 2015, Wood et al. 2015, Liang et al. 2017, Broadbent et al. 2018), (ii) reduced pain, anxiety and loneliness (Okita 2013, Robinson et al. 2013a), and (iii) psychophysiological improvement (Mitsui et al. 2001*a*, Saito et al. 2003, Wada & Shibata 2007b, Kawaguchi et al. 2012, Robinson et al. 2013c, Carretero & Sáez 2017, Petersen et al. 2017).

2.3.1 Improved social interaction

Paro has been used as a mediator for in-group or individual therapy in resident care homes and hospitals to improve social interaction between patients and staffs and relatives. Elderly people with dementia and autistic children may have some difficulties in their social skills such as problems with communication and lack of interaction with people. Elderly people with dementia can develop negative behaviours including verbal and physical aggression (Pedersen 2011, Moyle et al. 2016). The negative behaviour might be caused by depression triggered by difficulty in communicating, loneliness, and boredom (Wada et al. 2013, 2008). Autistic children also have social and learning difficulties. They also tend to develop repetitive behaviour and have lack of communicative abilities and imagination (Pipitpukdee et al. 2011).

Children:

Marti et al. (2005) investigated the use of the Paro robot as a mediator to stimulate social interaction in autistic young adults (n = 3). Participants' behaviour such as speech, eye movement, and behaviour when interacting with Paro i.e. patting, stroking, and holding were observed for three months. The researchers used storytelling and play-acting to observe the participant communication and interaction with therapist and robot. They found that the robot's physical characteristics i.e. facial appearance, shape, size, colour influenced the participants' attention, hence, increased their engagement during the social interaction.

Yamamoto & Kimura (2007) studied the behaviour of healthy children (n = 22) using video recording to understand how children behave when interacting with three different robots: AIBO, Paro and ifbot. The ifbot is a humanoid robot that has an ability to detect emotions based on listening to the voice tone and words when communicating with it. The ifbot robot has a few functions such as performing calculation, singing, giving advice, doing medical checks and it also includes a few games i.e. puzzle and memory games. However, the robot is not able to walk by itself. In the study, the robots were not introduced as robots. As a result, the children treated the robots as living animals. They also preferred to play with robots that were more interactive (ifbot) compared to AIBO and Paro. This is because compared to Paro and AIBO, only the ifbot robot can communicate with the children. The talking robot ifbot attracted the children to engage with it.

Pipitpukdee et al. (2011) also studied the ability of the robot to improve the behaviour of children with learning difficulties' behaviour when socializing with people. Children with learning difficulties can have problems to maintain eye contact and can exhibit repetitive behaviour. In the experiment, they used a toy Duck and a robot Paro and 37 participants aged 3 to 8 were recruited. Two conditions of therapy were applied: individual and group therapy. In individual therapy, observation showed that autistic children improved their eye contact with robot. However, in group therapy, autistic children tend to communicate with other children. Hence, the use of robot helps the children to improve their social interaction. The improvement of autistic children also reduces stress level among parents.

The latest study by Broadbent et al. (2018) compared the use of Paro and iRobiQ in a rural school. Participants aged between 2 to 18 years old (n = 207) had a 30 minutes session with the robot. Their interaction was observed and rated by an observer. Participants interacted with the robots in a group of 2 to 5 people. They found that participants were significantly more engaged and showed more positive expressions i.e. laugh with Paro compared to iRobiQ (a humanoid robot).

Young adult:

A study, by Wood et al. (2015) used a sample from the adult population (n = 114). The Paro was used to act as a mediator between two participants who were strangers to each other. The study was conducted in laboratory settings at University of Sheffield, United Kingdom. The conditions used two examples of Paro: active and inactive, and an active rubbery dinosaur robot, Pleo. Based on the results from questionnaires and interviews with participants, the study found that the active Paro was effective as a mediator to make conversation with strangers more enjoyable compared to an inactive Paro and Pleo.

Elderly people:

A study by Wada et al. (2004*b*) showed that there were only slight differences in mood improvements between groups of older people with dementia (n = 23) after interacting with an active Paro and an inactive Paro. Participants freely interacted with Paro in a group interaction. However, the presence of an active Paro increased laughing and communication between elderly people and staff in the nursing home.

Taggart et al. (2005) also investigated the behaviour of participants (n = 23) in nursing homes in response to an active Paro, an inactive Paro, or being alone with no robot as control conditions. In the no robot condition, participants tended to sit silently. But, in the switched on robot condition, they were more likely to actively interact with the robot and with the caregiver.

Kidd et al. (2006), studied the effect of Paro's reactive and proactive behaviour. Participants (n = 23) were introduced to reactive and proactive Paro by caregivers. The experiment ran for 4 months. Paro was placed on the table in public area. Everybody could play with Paro. Prior to the experiment, caregivers introduced Paro and they taught the elderly how to play with Paro. The results from video observations and Likert-Scale questionnaires showed that there were improvements in social activities on both conditions. The elderly people also preferred to play with a reactive Paro.

A study by Wada & Shibata (2007a), also showed that patients (n =12) tended to create conversations between other patients and staff after having been introduced to Paro and interacting with it. In the study, Paro was placed in public area where patients could play with Paro at the same time as interacting with other patients. They concluded that the presence of Paro triggered the intention to communicate with other patients. The results were also supported by de Graaf & Allouch (2014) and Sung et al. (2014), who found that engagement with an active Paro increased their communication and interaction skills in group activity.

The study by de Graaf & Allouch (2014) was adapted from a previous study by Kidd et al. (2006). The study was conducted in an elderly care home in Amsterdam. 17 elderly people were observed using video recording and participants' facial expressions were rated using a face scale. Then four elderly people were interviewed at the end of a session with Paro. The participants could freely interact with Paro in a group. However, the paper did not report the length of the session. Despite the positive finding during the group activity, however, the study also found that female participants were more actively engaged with Paro compared to male participants. The study conducted by Sung et al. (2014) observed a group of elderly people (n = 12) playing with Paro. The participants played with Paro in 30 minutes sessions two times a week for a month. The participants' social behaviour during the group activity were observed using Assessment of Communication and Interaction Skills (ACIS) and Activity Participation Scale. The study found a significant improvement in social behaviour from baseline to the final weeks of the session.

Takayanagi et al. (2014) also concluded that an active Paro robot can provide positive effects such as increased interaction between elderly people and caregivers compared to a stuffed toy. 30 participants with mild or severe dementia participated in this study. The study was conducted in participants' private rooms 15 minutes sessions over a three to six months period. The results from recorded facial expressions showed more negative expressions to the stuffed toy compared to when interacting Paro. It was suggested that the lack of responses from the toy decreased the positive experience.

A study by Yu et al. (2015) was based on elderly people with dementia (n = 40) among the Chinese population using two conditions of group activities: robot and table games. Elderly people with mild to severe dementia played with Paro for 30 minutes a day during a six-week period. It was found that the activities with the robot improved the moods and behaviours of the older people as compared to playing table games.

A recent study by Liang et al. (2017) also observed a significant improvement in communication between elderly people with dementia (n = 60) and caregivers after interacting with Paro compared to control condition. In the control condition, participants had regular activities such as playing words games or listening to music and dancing. Another significant improvement was observed in the facial expressions in Paro condition compared to control condition (no Paro). Participants played with Paro for one hour session for three times a week for 6 weeks at the dementia care centre and another 6 weeks at home.

2.3.2 Reduced pain, anxiety and loneliness

Other research has found that Paro can decrease loneliness, anxiety and stress, and improve moods (Okita 2013, Robinson et al. 2013a).

Children:

Okita (2013) suggests that Paro can be used in hospital for children to reduce pain and anxiety. The author focuses on how to develop coping skills using Paro as mediator. The patient played or interact with Paro in two conditions: (i) playing by themselves and (ii) accompanied by their parents. The result showed that levels of stress and pain decreased when the patient was accompanied by their parents. It was argued that this because Paro created a social bond between patients and their parents.

Elderly people:

A study by Robinson et al. (2013a) focused on the reduction in loneliness in elderly people. The elderly people were divided into two groups: one group (i) interacted with Paro and another group (ii) played with a resident dog. The loneliness of the older people was measured using a loneliness measurement scale. The elderly people interacted with Paro in a group by talking and patting Paro. It was found that elderly people tend to interact more with Paro compared to a resident dog and that there was a decrease in loneliness for the group who interacted with Paro.

2.3.3 Psychophysiological improvement

There are several studies that have been conducted that have shown positive outcomes of Paro on human psychophysiology (Mitsui et al. 2001*a*,*b*, Saito et al. 2003, Wada & Shibata 2007b, Kawaguchi et al. 2012, Robinson et al. 2013c, McGlynn et al. 2016, Carretero & Sáez 2017, Petersen et al. 2017, Crossman et al. 2018). In human-robot interaction, "psychophysiology" terms use in studies that investigating human emotions, behaviours and performance.

The effects of physical interaction with Paro have been measured using physiological measurements such as electrocardiogram (ECG), electrodermal response (EDR) and blood pressure (cardiovascular responses), respiration rate, and urine (stress hormones) and brain activity.

Children:

In a recent study by Crossman et al. (2018), the authors investigated the effects of short-term interaction with Paro on cortisol reactivity, anxiety (State Trait Anxiety Inventory (STAI)) and moods (Positive and Negative Affect Schedule (PANAS)) in healthy children. Participants ages between 6 and 9 (n = 87) were randomly allocated to one of three conditions: active Paro, inactive Paro and rest. They also were asked to perform a stressful task before the experimental condition session. The task was to tell a story and to complete a mental mathematical task and to be judged by two individuals in the room. The findings of the study found that there was an improvement in positive mood after interacting with Paro for 15 minutes. The authors suggested that the influence of Paro on moods showed that there was a potential that the robot can be used in situations to help reduce the effects of daily stress or depression in children. Unexpectedly, at the end of experiment, there was no reduction in cortisol reactivity as responses to stress in all conditions, when compared to baseline.

Young adult:

Research by Mitsui et al. (2001*a*) demonstrated the use of electrocardiogram (ECG), electrodermal response (EDR), and respiration rate to study human body changes while interacting with Paro by patting or stroking it. In order to evaluate the status of mood before and after interacting with Paro, the data collected using the physiological measurements were analysed together with Profile of mood state (POMS). In the experiment, they investigated the differences in physiological changes in four different interactive sessions that involved (1) an object that had similar fur and dimensions to Paro and (2) a column made of vinyl and playing with Paro when it was (3) switched on or (4) off. The study identified elements that contributed to the mood changes. They found that there were increases in heart rate variability (HRV), respiration and EDR when participants were interacting with the active Paro compared to other conditions. The heart rate (HR) did not increase in the active condition, but the HR did increase when participants stroked a column covered with Paro's fur and an inactive Paro. However, the result showed no major differences in the different cases tested because the data sample was very small (n = 6).

Later, in a study by Kawaguchi et al. (2012), the authors aimed to understand the effects of interacting with active or inactive Paro on the brain. The study used functional near-infrared spectroscopy (FNRIS) to collect the physiological data and participants were asked to perform a mental task before interacting with the Paro. The mental task was required to balance the brain activity before interacting with Paro. The participants were given the mental task to solve and then asked to play with Paro ON or OFF for 1 minute. After they played with Paro, the result showed that there was an activation of brain activity while Paro was turned on or off. But, interestingly, when participants interacted with active Paro, there was activation in an area of the brain related to emotion recognition and efforts to communicate. Then, when participants talked to inactive Paro is may related to speech. The results suggested that the interaction with active Paro not only create an effort to talk but also induced emotional responses towards active Paro. However, similar to previous study of Mitsui et al. (2001*a*), there are some limitations to the Kawaguchi et al. (2012) study: there was no comparison with a control group and the number of participants was really small (n = 10).

Elderly people:
Saito et al. (2003) and Wada & Shibata (2007b) used urine testing as a physiological way to measure the effect of Paro on stress hormones. They studied the hormone in the urine named 17-ketosteroid sulphate (17-KS-S) and 17-hydroxycorticosteroisa (17-OHCS).

In 2003, Saito, Shibata, Wada & Tanie investigated changes in stress hormones as the result of using Paro among elderly people living in nursing home (n = 23) and the caregivers (n = 30) by looking at levels of 17-KS-S and 17-OHCS. In the experiment, there were two conditions of Paro, the proactive and reactive behaviour. The proactive Paro can exhibit most of Paro's behaviour such as blinking its eyes, swinging its fins, moving its head, and crying for attention. However, the reactive Paro has limited behaviour with only ability to cry and moving their head. Participants took turns to play with Paro in the group. The length of the study was three weeks and participants got a chance to interact with Paro four days a week. Participants and caregivers' urine was taken on the first week of interaction with Paro and at the end of every week. Caregivers did not participate in playing with Paro during the therapy session but they were present in the room. Based on urine analysis, it was suggested that levels of stress decreased after the elderly people interacted with proactive Paro. Caregivers' level of stress were also reduced after Paro had been introduced to elderly people because the elderly people no longer needed extra attention from the caregivers.

Wada et al. (2005b) used electroencephalography (EEG) to measure cortical neuron activity (nerve cells).and the data collected was analysed using the Diagnosis Method of Neuronal Dysfunction. The data was collected before and after interacting with Paro. The setting of the study was in a Japanese care home with a free group interaction in a 20 minutes session. Participants were elderly people with mild to severe dementia (n = 14). The results found an improvement in cortical neuron activity in the elderly people.

Wada et al. (2005b) and Kawaguchi et al. (2012) studied the effects of Paro on brain activity using electroencephalography (EEG) and functional near-infrared spectroscopy (FNRIS) respectively. Both of these measurements measure changes in mental states. However, the EEG measures the electricity triggered by neuronal activity. FNRIS measure the changes in blood volume and oxygenation as the result of homo dynamic responses to activity in the brain (Hirshfield et al. 2009).

In 2007, Wada & Shibata used interviews, urine analysis, video, and case study to analyse the improvement of social interaction in people who lived in a Nursing Home (n = 12). The urine was taken before and after interaction with Paro over three weeks: (i) a week before interaction with Paro (ii) 2 weeks after and (iii) week 3. Before the introduction of Paro, some elderly people spent more time together talking and playing cards. However, some other elderly people spent less time socializing with each other and just walked past in the corridor. After the introduction of Paro, the time spent socialising was increased. The stress hormones in the urine also showed a positive improvement with an increasing number of 17-KS-S and 17-OHCS ratio value.

A study by Robinson et al. (2013c) measured blood pressure and heart rate to find out the physiological effect of Paro using cardiovascular measures. The study was conducted at a retirement house in New Zealand. The experiments were conducted for 12 weeks. The aim of the study was to see the changes in heart rate and blood pressure as the result of interaction between Paro and elderly people (n = 17). In the experiment, the blood pressure of participants was taken before and after the interaction with Paro. Before collecting their physiological data, the participants were left alone resting in the room for 5 minutes. There were significant decreases in blood pressure (p = 0.048) and heart rate (p = 0.04) in participants who engaged in physical interaction such as stroking and patting the Paro.

In 2016, McGlynn, Geiskkovitch, Mitzner & Rogers investigated the impact on stress when either an active or inactive Paro was present or not present during stressful situation. In order to induce stress, a Raven's Progressive Matrices Test (RPM) was used, and to assess levels of stress State Trait Anxiety Inventory (STAI) was used. Participants were assigned to one of two conditions, then they were asked to answer a set of matrices intelligence quotient questions (RPM) in 5 minutes with or without the presence of active or inactive Paro. Participant were also asked to rest for 3 minutes with the presence of active or inactive Paro. The authors found an increased in stress during the stressful task when Paro not present. But, the authors also found a similar trend of stress reduction after the stressful task in both the active and inactive Paro conditions.

A recent study by Carretero & Sáez (2017) found a significant reduction in pulse pressure but not in heart rate in people with Alzheimer's disease or Dementia (n = 11) after interacting with Paro for 20 minutes each session for four consecutive weeks.

Another recent study of Paro by Petersen et al. (2017) found a reduction in stress and anxiety according to physiological measures (i.e. pulse rate, galvanic skin responses) after interacting with Paro for a few weeks. The effectiveness of Paro intervention was compared to a control condition. The participants were divided into two groups: the first group had a 20 minutes session with Paro 3 times a week, the second group also had a 20 minutes session that include physical activity, listening to music and doing mental activity. The study had 61 elderly people with dementia and the experiment was conducted at a dementia unit in United States. These results suggested an improvement in oxygenation and cardiovascular status in a 3-month intervention program.

As mentioned in the previous section 2.3.1, a study by Liang et al. (2017) not only observed the interaction between elderly people and caregivers with Paro as mediator, but, also measured the changes in blood pressure, salivary and hair cortisol. The study compared between two conditions: (i) Paro and (ii) standard care in care home. At the dementia care centre, the blood pressure and salivary cortisol were collected from the elderly people before and after each session. Then in home settings, blood pressures were collected at beginning of the study, and after 6 weeks and 12 weeks of interventions. The hair cortisol, however was collected at baseline and after 6 weeks' period. Unexpectedly, the results found no difference in physiological responses between both of the conditions.

A summary of previous studies of Paro on psychophysiological improvement can be seen in Table 2.2.

Table 2.2: Previous studies of Paro that looking at physiological measurements

Physiological measurements	Country	Participant	Method	Result	Author
EEG, Respiration rate, EDR	Japan	6 participants, 25-30 years, healthy	POMS	Positive mood changes	Mitsui et al. (2001*a*)
Urine (17-KS-S and 17-OHCS)	Japan	60 participants, healthy and light, heavy dementia	Interview scaled of Holmes and Rahe, questionnaire, Wilcoxon rank, SPSS, Burnout score	Improvement in level of 17-KS-S and 17-OHCS	Saito et al. (2003)
EEG	Japan	14 participants, elderly, mild and severe dementia	DIMENSION	Improve mood	Wada et al. (2005a)
Urine (17-KS-S and 17-OHCS)	Japan	12 participants, 67-89 years, Healthy	Case study	Improvement in level of 17-KS-S and 17-OHCS	Wada & Shibata (2007b)

Continued on next page

Table 2.2 – continued from previous page

Physiological measurements	Country	Participant	Method	Result	Author
fNRIS	Japan	10 participants, 21-33 years, Healthy	Word association, 35 channels of brain, Paired sample t-test	Brain activation in speech and brain area	Kawaguchi et al. (2012)
Blood pressure, heart rate (CardioScope II)	New Zealand	21 participants, 71-95 years	Long-term interaction	Blood pressure decreased	Robinson et al. (2013c)
Pulse and heart rate	–	11 elderly with Alzheimer's disease and dementia	Long-term interaction	Reduction in pulse pressure	Carretero & Sáez (2017)
Pulse	United States	61 elderly people with dementia	Long-term interaction in a group	A reduction in stress and anxiety according to pulse rate	Petersen et al. (2017)
Saliva and blood pressure	New Zealand	30 dyads, pair of elderly people with dementia and their caregiver	Video observation and long-term interaction	Improve social behaviour of elderly people with the caregiver. But, there was no change in physiological responses between both conditions	Liang et al. (2017)
Salivary cortisol	United State	87 children ages between 6 to 9 years old	PANAS and STAI	Improve positive moods. No improvement in cortisol.	Crossman et al. (2018)

2.4 Stress and stressors

This section reviews about stress and stressor that can be used in short-term stress reducing intervention in laboratory settings.

What is stress?

Stress is defined as a mental state that affects several physiological and psychological functions of humans such as increases in resting heart rate and blood pressure which are related to emotional changes (Taelman et al. 2009, von Dawans et al. 2012). Everybody can suffer from stress regardless of age and gender. Stress can happen any time and everywhere. The symptoms of stress can be frequent headaches and various health issues such as gastric problems, anxiety, irregular blood pressure, and heart rate. Surprisingly, people might not be aware that they have these symptoms (Spielberger & Sarason 2013).

Types of stress

There are three types of stress: (1) Acute, (2) episodic and (3) chronic stress. Firstly, acute stress is an immediate reaction to certain tasks such as public speaking and taking examinations. The average time for coping with stressful events depends on individual variabilities (e.g. age, gender, lifestyle). After acute stress, the body will relax and return to resting levels. Another type of stress is episodic stress. Episodic stress is a series of acute stress episodes. For example, when one stressful event occurs and is followed by other stressful events (Roberts 2011). People who suffer from episodic stress will become fatigued due to negative emotions triggered by stressful events and if the stress is left untreated over a long period, that person may develop chronic stress. Chronic stress is a prolonged form of episodic stress that was not properly managed (Sahoo & Sethi 2015). Chronic stress can lead to serious health issues such as a weak immune system, low metabolism and cardiovascular diseases (Choi et al. 2014, Mah et al. 2016). Prolonged stress also increases the possibility of harming the self and others (e.g. suicide and domestic violence such as abuse and homicide) (Saito et al. 2003).

The relationship between levels of stress and stressors can be different based on people's age, gender, environment, and lifestyle (Fukuda et al. 2001, Cacioppo et al. 2007, Janipha et al. 2012). For example, a working adult may be stressed out by workload, family responsibilities and financial concerns (Leger 1996). However, for students, an academic workload (i.e. learning, homework, examination, class assessment) are the main causes. Another factor for student stress can be loneliness (e.g. homesickness) (Sawir et al. 2008). Lifestyles of heavy drinking, smoking, and poor daily nutrition intake can also increase levels of stress (Fukuda et al. 2001). Other external stressors are related to the environment such as air and sound pollution produced by transportation and factories (Cacioppo et al. 2007).

Since stress can affect people at any age, based on stress surveys in younger and older generation, it has been reported that the younger generation has higher levels of stress compared to older generations (Andrews & Wilding 2004, Shields 2004, Kennedy 2015, Coccia & Darling 2016). The young generation age group is between 18 to 40 years old. In the young age group of younger populations, it has been shown that students were experiencing higher levels of stress, depression and anxiety (Petry 2002, Dyrbye et al. 2006).

Stress among students

Previous studies have reported some risk factors related to stress among students. A study found that among students, females are more likely to have higher levels of anxiety and depression (Rose & Rudolph 2006). Another risk factor that contributes to student stress is poor time management. Due to academic demands, students do not manage their study and leisure time properly and wisely. The academic workload can also disturb sleep patterns and affect the student levels of stress (Lund et al. 2010). Increasing levels of stress are experienced by students during academic semesters (Misra & McKean 2000). Other studies reported that students who are studying medical, dentistry, and engineering show higher levels of stress (Kumaraswamy 2013, Waghachavare et al. 2013).

Family background can also be related to stress among students. Previous studies reported that students' anxiety and stress behaviour were associated with their family financial situation (Misra & McKean 2000, Andrews & Wilding 2004, Eisenberg et al. 2007).

Apart from that, one of the causes of students' stress is noise. Students who stay in shared accommodation tend to show higher depression, anxiety and stress because of noise, house condition, problems with cleanliness of other occupants and bad relationships with house mates

(Heath & Kenyon 2001, Andrews & Wilding 2004). Other risk factors that increases levels of stress are bad eating behaviour (i.e. eating junk food), poor life-style (i.e. lack of exercise) and abnormal sleep behaviour (Hudd et al. 2000, Lumley & Provenzano 2003, Brougham et al. 2009).

Stress induction

In order to measure the stress responses, especially in laboratory settings, task induced stress is used to induce acute or short-term stress. There are some different techniques that can be used for stress induction in laboratory settings, in order to induce immediate changes in human physiology and psychology (Dickerson & Kemeny 2004). For example, stressful tasks that use mental workloads such as interview, examination, public speaking and playing video games, Montreal Imaging Task (MIT) (which includes mathematical task), and watching a stressful video, listening to a noise, and physical stressor (such as cold pressure and electric shock) (Dedovic et al. 2005, Beetz et al. 2012, Lass-Hennemann et al. 2014, Luijcks et al. 2014, Nguyen & Zeng 2014).

There are three types of stress induction: (i) cognitive tasks, (ii) cognitive and verbal public speaking (e.g. Trier Social Stress Test (TSST), and (iii) noise exposure (Dickerson & Kemeny 2004). Dickerson & Kemeny (2004) mentioned that verbal public speaking could be an effective stress induction to create acute stress in the laboratory and create a higher level of cortisol activation. Cortisol is a hormone that is related to stress responses. However, another study showed that there is activation of heart rate and skin conductance while performing laboratory stress tasks with noise and electric shock (Luijcks et al. 2014). The activation of heart rate and skin conductance are also related to stress responses and mood changes. Stress induction will influence mood changes from a restful condition to negative moods (Luijcks et al. 2014).

Previous studies on stress reactivity used the TSST protocol for children as a stressor to examine levels of stress (Beetz et al. 2012). The children were asked to perform stressful tasks such as doing presentations in front of people and mathematical tasks. The levels of stress were measured before and after the stressful tasks. The measurements consisted of salivary cortisol, questionnaires: Separation Anxiety Test (SAT) and Self-Assessment Manikin (SAM), and behavioural recording video. In their findings, it was found that there was an increase in levels of cortisol after the stressful tasks (Beetz et al. 2012).

In another study by Lass-Hennemann et al. (2014), they used a traumatic video to induce stress. The authors measured the subjective evaluation using Positive and Negative Affect Schedule (PANAS) and State Trait Anxiety Inventory (STAI). At the same time, levels of stress experienced during and after the stressful task were measured using cortisol, heart rate, and blood pressure. Lass-Hennemann et al. (2014) found an increase in blood pressure, heart rate and cortisol after watching the traumatic video.

Noise as a stressor

Another interesting tool to induce stress is noise. Noise is known as one of environmental stressors (Westman & Walters 1981) that can negatively impact mental (Stansfeld & Matheson 2003) and physical health (Basner et al. 2014).

The noise can be a suitable tool to induce stress in laboratory experiment because it can induce mild stress in a short time (Mahmood et al. 2006). Noise can produce negative effects in humans such as annoyance, frustration, and excitement, anxiety, anger, distortion, disappointment, dissatisfaction and helplessness (Westman & Walters 1981, Stansfeld & Matheson 2003, Goines & Hagler 2007). The levels of acceptable decibel (dB) sound that can be heard by human are around 35 to 45 decibel Hsu et al. (2010). However, unwanted or unpleasant sounds called noise are those higher than 55 decibels (Kumar et al. 2004).

Noise has effects on human hearing, sleep, task performance, communication, and emotional states (Westman & Walters 1981). Noise also induces short-term or acute stress. The short-term or acute stress induced by noises will temporarily increase blood pressure, heart rate and affect moods. Increased heart rate is a response that indicates readiness to fight or flight when being exposed to loud or unfamiliar sound. The sound might trigger different reactions in people who hear it, it depends on individual judgements about the sound. For example, some people will determine unwanted sound such as loud music and road traffic as noises but some people find the sound is not disturbing (Rylander 2006). Goines & Hagler (2007) also mentioned that the behaviour of people who are being exposed to noise levels of more than 80 decibels can be aggressive and may reduce their helpfulness to others. The noise also increases levels of blood pressure and heart rate. These increasing levels are related to the activation of the neuroendocrine system and the activities of sympathetic system in the human body (Mahmood et al. 2006, Wang et al. 2007).

Noise induction has been shown to increase stress levels in people exposed to different levels of noise and uncontrollable noise (Sim et al. 2015). The combination of low and high noises may induce stress. Mahmood et al. (2006) used short-term noise exposure with continuous similar frequency and intensity. They reported that there are increasing levels of blood pressure and heart rate between 2 to 11 minutes after being exposed to noise at 90 decibels. The levels of heart rate and blood pressure after exposure to the noise remained higher compared to the resting conditions. It was found that participants take several minutes to return to the rest conditions. The average time taken by the heart rate to come down to the rest condition level is 2 minutes.

Another study by Geen & McCown (1984) investigated the effects of controllable and uncontrollable noise on skin conductance, heart rate, and blood pressure. They mentioned that an uncontrollable condition would increase stress related symptoms because participants would feel more uncomfortable with uncontrollable stimulation. For example, during controllable conditions, the participants can control the noise stimulation during the experiment. Therefore, they will feel less stressful compare to the uncontrollable condition. However, Kumar et al. (2004) reported that participants exposed to an increasing level of continuous white noise from 65 to 90 decibels showed an increase in their skin conductance level.

2.5 Physiology of stress

It is important to understand the nature of stress involving producing and controlling stress responses. The stress responses include increased heart rate, blood pressure, sweating and emotional reactions or changes.

The psychophysiology of stress refers to the changes of physiology in our body when it reacts to the stressor and the reaction to stress is related to our mind. The effect of the stressor on the autonomic nervous system (ANS) is received by the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The SNS will decide to fight or flight in response to a stressor. However, the PNS will act as the relaxation response after a physical or psychological reaction to a stressor (Taelman et al. 2009).

The flight and fight response was described by Selye (1974) in the General adaptation syndrome (GAS) of stress (Goldstein 2010, Szabo et al. 2012). There are three fundamentals of GAS which concern the activation of the sympathetic nervous system to react to the stress. In the second stage, the body will resist the stressor, however, the time to cope with the stress depends on the individual. The examples of reactions to the stressor are physiological arousal or mood changes. Third and finally, if the body is exhausted, and no longer able to cope with the stressor, it will lead to various health issues such as heart attack, diabetes, and impaired immune system (Vanitha & Suresh 2014).

Based on GAS of stress, the flight and fight response by SNS can be the negative emotions and physical reactions such as facial expressions, changes in pupil size, increase in heart rate and blood pressure, sweat glands activation and increased respiration rate (Dang & Tapus 2013, Lass-Hennemann et al. 2014). When stress triggers, neurochemicals such as epinephrine (adrenaline), norepinephrine (noradrenaline), acetylcholine, and cortisol are released by the adrenal glands as a response to the stressor. These increase the activity in the respiration and cardiovascular systems. The activation of SNS will also increase the activation of sweat glands.

The sweat glands are related to the other parts of the brain such as the amygdala (which functions to decode emotions and act as stimuli to any harmful event to our body), the hippocampus (related to memory and emotions), and the prefrontal cortex (controlling anxiety). The affective processes such as memory, attention and emotional responses occur in the limbic system. Additionally, both SNS and PNS cannot act simultaneously, therefore human cannot feel relaxed and stressed at the same time (Seaward 2012).

Then in order to know more about how to assess the human responses to stress or to Paro, the following sections review various methods that have been used to evaluate the human responses or emotions.

2.6 Assessing human emotions

Based on the previous section 2.3 on page 16, the Paro robot has been found to have positive effects on emotions and physiological responses during social interaction. Various methods have been used to evaluate human emotions, such as self-reports, interviewing, physiological responses, task performance metrics and video behavioural observation (Bethel & Murphy 2010). However, the most common methods used to assess human emotion are self-reports and physiological responses. In this thesis, one goal is to assess whether interaction with Paro can reduce stress responses using self-reported, video observation and physiological responses. Therefore, the following sections review the fundamental aspects of human emotions and how to assess the emotions.

2.6.1 Emotions and facial emotional expression

Human emotions can be defined as emotional states or reactions that can be bodily based responses (Berne 1972). Based on the Oxford Dictionary, emotion defines as *"a strong feeling"* resulting from a certain event and moods (Oxford English Dictionary 2017). There are several aspects of humans that are influenced by emotions including as cognition or memory, physiology, expression and feeling (Oatley & Jenkins 1992, Kolling et al. 2016).

Emotions are associated with short term durations, but, moods can last for hours and days (Lottridge et al. 2011). There are six basic types of emotions: anger, happiness, sadness, fear, disgust and surprise. However, there are some arguments about the numbers of basic emotions, the numbers could be four, five, six or eight (Ekman 1992, Lottridge et al. 2011). Then Jack et al. (2016) suggested that there are only four basic emotions, due to overlaps on perception between surprise and fear emotions, and between disgust and anger emotions. Other emotions that can be categorised as non-basic emotions are boredom, relaxation or neutral, excitement, and confusion (D'Mello & Calvo 2013).

Based on theories of emotion, there are two categories of emotion i.e. basic and dimensional. The basic theory has been explained above. However, the dimensional theory is a combination of emotions in two dimensions. The dimensions can be categorised in arousal and valence (Russell 1980). Arousal can be defined as the activation of emotions that are stimulated by the automatic nervous system. This activation of emotions is related to physiological changes in skin conductance, heart rate and other physiological responses. The physiological changes caused by a specific event (e.g. stress) will influence emotional states. The arousal can range from low (i.e. calming or soothing) to high (i.e. excitement or agitated). The arousal can be measured using self-reported and physiological measurements. Valence, however, provides a clue to positive or negative emotional states (Lang 1995). These two fundamental dimensions, arousal and valence were introduced in a circumplex model of emotion by Russell (1980).

Emotions can be identified using facial expressions, psychological and physiological measurements. The psychological and physiological measurements will be discussed in next section 2.6.2 and 2.6.3. Early research on facial expression and recognising emotions was conducted by Darwin in 1872. Darwin (1872) mentioned that the facial expressions of emotion are recognised similarly around the world. However, there are some arguments that the perception of emotions can be different in different cultures (Jack et al. 2012).

Facial expressions are defined as muscle movement beneath the skin on the face (Ekman 1993). Examples of muscle movement are frowning, smiling, eyes' widening and nose wrinkling (Ekman et al. 2002). The basic emotions such as anger and happiness are easier to distinguish and identify visually (Silva & Miyasato 1997). For example, happiness can be identified by looking at smiling with the corner of the lip pulled up and backwards, the cheek lifted and visible wrinkles around the eyes (Ekman et al. 2002). Happiness can be related to pleasure and excitement experiences (Ekman 2003). It can be defined as the feeling of achieving something (Dursun et al. 2010). It includes smiling and laughing. Disgust can be defined as feeling of revulsion or disapproval of anything that is extremely unpleasant (Ekman 2003). Feelings of disgust can be caused by touching, looking, smelling or tasting unpleasant things or by just listening to an idea (Ekman 2003).

Then, example of a non-basic emotion is boredom. Boredom is defined as dissatisfaction due to lack of interest and activity and can also be related to low arousal based on the dimensional theory (Kroes 2005). Boredom is categorised as negative experience (Fahlman et al. 2013). However, there are mixed arguments about the physiological changes that occur during the experience of boredom. It has been claimed that, boredom can increase heart rate and skin conductance (London et al. 1972). But another study suggests that boredom is associated with both high and low arousals (Eastwood et al. 2012). Bench & Lench (2013) proposed that boredom is the state of an emotion that changes from a higher intensity of emotion to one approaching a neutral state. This can lead to mind wandering which is related to disengagement with current task (Smallwood et al. 2007). The expression of boredom can be identified by closed lips, open eyes with slightly closed eyes' lids, blank gaze, and sometimes a person who experiences boredom will roll their eyes (Bench & Lench 2013). Boredom is often confused with sadness (Silvia & Warburton 2006).

Another example of a non-basic emotion is confusion. Based on the Cambridge dictionary, being confused is defined as *"(of a person) unable to think clearly or to understand something."* It could be because of *"a lack of clear instruction or not enough information so difficult to understand"* (Cambridge 2015). The facial expression when someone is confused is defined by wrinkled nose and crease between the eyebrows (FaceTurn 2011). However, the facial expression can be easily misinterpreted, for example, there are confusions when identifying emotions between fear and surprise, between disgust and anger, between sadness and disgust, between sadness and neutral expressions (Palermo & Coltheart 2004, Calvo & Nummenmaa 2008).

Overall, the facial expression does not only help to tell the current emotional state of individual, but the face is also has been significantly important in social interaction as form of a non-verbal communication (Ambady & Rosenthal 1992, Ekman 1992, Öhman 2002). Facial expressions allow a first person to express their internal feelings to a second person in social interaction (Calvo & Nummenmaa 2016).

2.6.2 Self-reported responses

Self-reported responses can be used to access the participant's personal feelings or opinions or emotional states (Shibata et al. 2003). The methods include questionnaires, surveys, and psychometric scales. However, a limitation of this methods is that participants might ignore their own feelings or thoughts when answering the questionnaire and try to answer the questions based on the researcher's view (Bethel & Murphy 2010). Other limitations of questionnaires are that participants might lose their concentration and participants might affect their mood if the questions are too long (Travis 2002).

There are many questionnaires that can be used to measure stress responses such as: State Trait Anxiety Inventory (STAI) Y-Form, University of Wales Institute of Science and Technology (UWIST) Mood Adjective Checklist (UMACL), Profile of mood state (POMS), and Positive and Negative Affect Schedule (PANAS). STAI Y-Form by Spielberger (1987) is used to measure levels of anxiety, and the UWIST UMACL by Matthews et al. (1990) used to measure changes in general arousal, energy, tension and hedonistic. POMS by Terry et al. (2003) and PANAS can be used to measure current feelings and moods.

Other self-report ways to measure prolonged acute stress include Stress Response Inventory (SRI) to measure tension, aggression, somatization, anger, and depression, fatigue, and frustration (Koh et al. 2001). The SRI is suitable for measuring severe stress states (e.g. depression). Perceived Stress Scale (PSS) by Cohen et al. (1983) is used to assess to what extent individuals perceive their life as stressful.

2.6.3 Physiological measurements

Physiological measurements can be used to measure any physiological changes in the human body. The physiological measurements have gained popularity among human robot interaction researchers, in order to measure human experiences when using the technology. Some examples of physiological measurements include skin conductance, cardiovascular, respiration, brain activity, muscles, peripheral nervous system, and others (Bassett et al. 1987, Ward & Marsden 2003, Hjortskov et al. 2004, Ghaddar et al. 2013).

The physiological responses can be affected by factors and events, which contribute to any changes in physiological responses. The factors and events can sometime be called as physical and mental stimuli. One example of physical stimuli is cold pressure task, where participants were asked to put their hand in the cold water for a minute (von Baeyer et al. 2005). One example of mental stimuli is a task. Participant can be asked to perform a different task with different levels of difficulties. Commonly, the task involves decision making, problem solving and mental imagery.

The changes in physiological responses can be either positive or negative responses or feedbacks (Ward & Marsden 2003). The physiological measurements provide a graph or report that gives clues to the physiological changes during the performance of activities and can enable identification of the significance of the events (Dawson et al. 2007). The physiological measurements also can be used to identify human emotions such as anger, grief, and sadness (Mandryk & Inkpen 2004).

Some examples of the use of physiological measurements, as follows: Figner & Murphy (2011) reported that skin conductance (SC) could reliably measure stress and affective states such as emotion and experience during certain events. Then Henelius et al. (2009) stated that heart rate (HR) can provide information on mental workload. Cowley et al. (2016), Shaffer & Ginsberg (2017) stated that electrocardiogram (ECG) can be used to monitor user concentration and engagement with a given task, based on changer in heart rate variability (HRV). At the same time, the ECG is able to monitor the ability of an individual to control their physiological reactivity as a response to stressful situation (Shearer et al. 2016).

According to Dirican & Göktürk (2011), physiological measurements have great potential to provide a better understanding of human responses. However, there are a few advantages and limitations of physiological measurements highlighted by the authors. The advantages of using physiological measurements that the data do not depend on human opinions. Therefore, participant cannot easily manipulate their responses. This helps to reduce self-reported bias (Arkin & Moshkina 2014). The measurements can be used together with self-reported responses and recorded video to support the reliability and validity of the data or to give additional information about the human reaction. Captured video is also very helpful as a time reference for giving the clues about the movement or intention of the participants during the experiment (Bethel & Murphy 2010, Tiberio et al. 2013).

Physiological measurements have sensitive sensors to sense any physiological changes that occur in human body during real time monitoring (Dehais et al. 2011). The sensor will capture different levels of cognitive and affective changes in both short time and long time (Dehais et al. 2011).

Despite the advantages, according to Dirican & Göktürk (2011), the physiological measurements can have some limitations in terms of cost and data collection. The cost of equipment can be expensive and the researcher needs to undergo training to use the equipment correctly. It is necessary to wear the equipment correctly to avoid noise in the data set. Another disadvantage of using physiological measurement, wearing the equipment might limit the participants' movement because the participants' body will be attached to the equipment and this could lead to unnatural behaviour. Therefore, data pre-processing is needed to reduce the data noise. Physiological measurements also produce a large volume of data, therefore, a researcher can easily misinterpret the meaning of data.

Other concerns:

In order to study physiological measurements, it is crucial to consider the types and sources of human variability such as age, gender, ethnicity, diets, lifestyles (e.g. smoking, alcohol consumption, physical activities, medication intake and body weight and height) (Sim et al. 2015, Picard & Wolf 2015). Physiological measurements are also not able to distinguish between activation such as anger or joy, both emotions will activate high levels of arousal. However, the results from physiological measures can be supported by using different measures (i.e. self-reports) (Arkin & Moshkina 2014).

The following section presents three commonly used physiological measurements in physiological research: (i) skin conductance, (ii) heart rate and (iii) heart rate variability.

2.6.4 Skin conductance

Galvanic skin response (GSR) is defined as any variation in skin conductance. Historically, the term of GSR was used in 1889, and then in 80s, the commonly used terms are electrodermal response (EDR) or skin conductance (SC) (Pour et al. 2010, Fish et al. 2018). According to Cowley et al. (2016), skin conductance can be reliable to measure human experience in human-computer interaction, as the measurement is easy to install and use. At the same time, the data collected form the measurement also easy to interpret. Furthermore, Cowley et al. (2016) suggested that the used of skin conductance is suitable as an introduction to physiological measurement. The EDA can provide a clue to sympathetic nervous system activity based on arousal (Dawson et al. 2007, Iffland et al. 2014).

The skin conductance is related to sweat production and used to measure anxiety, fear, joy, and stress levels (Dawson et al. 2007, Tiberio et al. 2013, Iffland et al. 2014). Sweat glands activities produce during fight or flight activity that occurred during stress responses. The sympathetic nervous system is part of the autonomic nervous system that is responsible for sweat production and it is related to affective processes. The areas of eccrine sweat glands are mainly on hands and fingers.

According to Westeyn et al. (2007), skin conductance measures the changes across two regions of the skin in electrical resistance. The change in conductivity in arousal states of individuals can be different based on differences in human mental, physical, and emotional responses. The states of arousal can be related to the activation of the sympathetic branch in the autonomic nervous system that is related to affective processes (Figner & Murphy 2011). The eccrine glands in skin will produce ionic sweats, which, lower the skin resistance and increased the conductivity during sympathetic responses. The measurement for skin conductance is in micro Siemens (µS).

There are two types of skin conductance responses: (i) tonic and (ii) phasic. The (i) tonic conductance (slow adaptation) is the baseline level of skin conductance that slowly changes over time and it can be different for each person because the level of arousal depends on each person's psychological state. The (ii) phasic skin conductance (fast adaptation) is associated with changes in human skin during short-term events and the changes influence by discrete environmental such as sound, light and smell (Braithwaite et al. 2013).

The physiological measurement can measure basic emotions such as fear, anger and sadness. Skin conductance is correlated with level of arousal and valence that related to positive and negative emotional responses (Pour et al. 2010). In Figure 2.2 shows the valence and arousal for the emotional states.

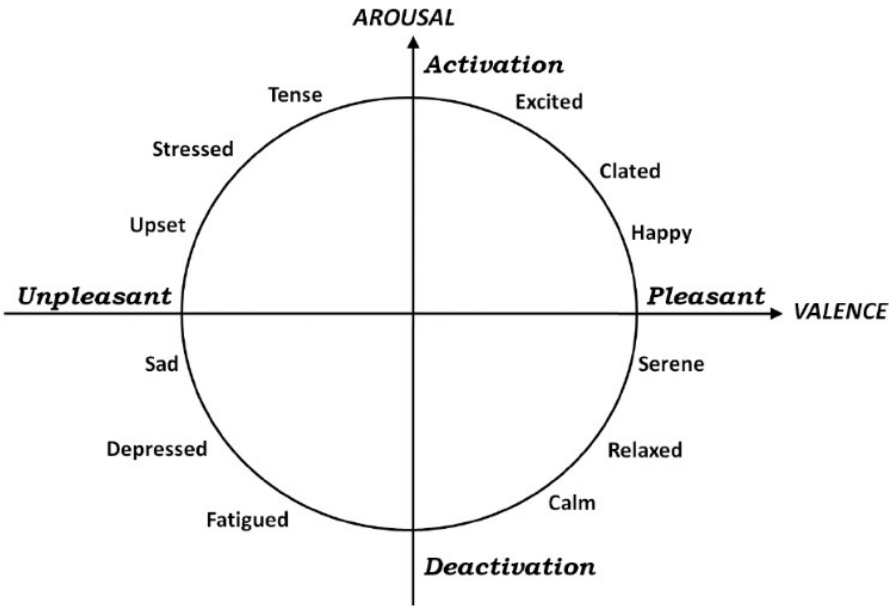


Figure 2.2: Arousal and valance classes (Adapted from Posner et al. (2005))

Despite the benefits of skin conductance in measuring human responses, the skin conductance is also very sensitive to a little movement and as a consequence movements can affect the data. Therefore, when collecting the data from the skin conductance, participants will be advised to avoid any major movement.

Data recording

In laboratory settings, the skin conductance can be recorded using a BIOPAC system (BIOPAC 2017). BIOPAC system includes hardware and software for collecting and analysis physiological data (Pflanzer 2005). Skin conductance sensors will be attached to the participants' non-dominant hand. In Figure 2.3 shows the participant wearing the skin conductance sensors by placing the two electrodes in a plastic bar on the middle finger and index. Participants need to maintain the pressure on the electrodes because if there is major variation in pressure will affect the results (Calvo et al. 2009).

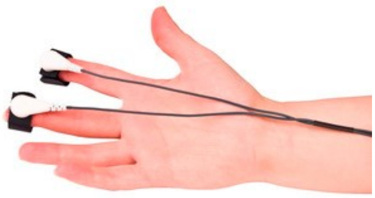


Figure 2.3: The skin conductance sensors were attached to participant's fingers (from Calvo et al. (2009))

In accordance with common recommendations by Cacioppo et al. (2007) and Figner & Murphy (2011), for data acquisition parameters, the low-pass filter is set to 1 Hz (for reducing noise in the collected data). Then the amplification of the skin conductance can be set to 0.01 to 5.00 µS as a boundary to physiological responses (Venables & Mitchell 1996, Kushki et al. 2011). Any amplitudes below the 0.01 or above the 5.00 µS will be considered as unrelated events. The term of "unrelated events" is defined as when there was a presence of amplitudes or peaks in SCR that occurred without any presence of physical or emotional stimuli (Braithwaite et al. 2013). The relevant to set the range of amplitudes as some people might have abnormal arousal that related to over reactivity of autonomic nervous system (ANS) (Lajante et al. 2012), and the sensor of SCR can be sensitive to any artefact movements (Boucsein et al. 2012). Furthermore, the choices of amplitudes can be influenced by the types of experimental conditions.

The data can be sampled between 100 to 2000 Hz (Braithwaite et al. 2013). However, a large data sample will require a large computer storage. Figure 2.4 shows the data recorded using BIOPAC MP36 data acquisition unit.

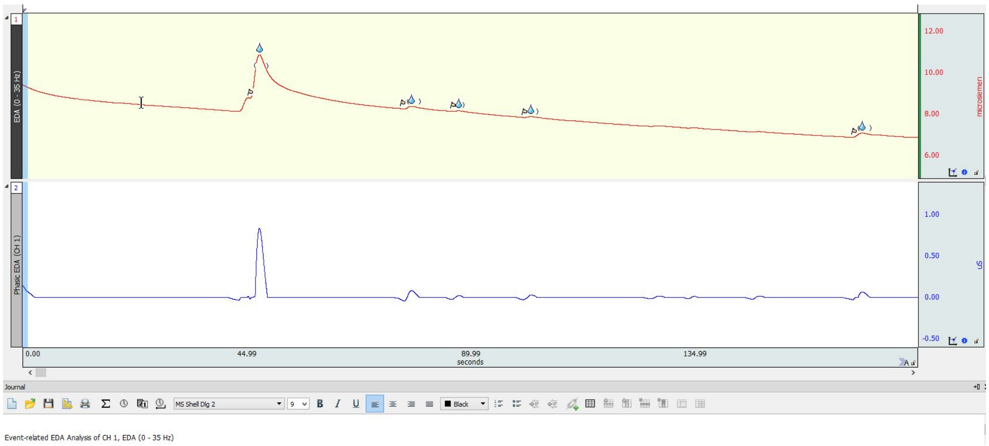


Figure 2.4: Example of skin conductance data collected from BIOPAC

Data analysis

In accordance to recommendation by Figner & Murphy (2011), Boucsein et al. (2012) and Braithwaite et al. (2013), the physiological data collected can be analysed using the BIOPAC software. One of most common data analysis used is calculating the event-related based on numbers of successful peak. As mentioned earlier, the event-related was count based on the amplitudes set, which between 0.01 to 5.00 micro Siemens µS. Figure 2.5 shows an example of the output of SCR from Excel. Based on the figure, the results showed that there was five events-related that has been detected.

Stimulus Matching Summary					
Amplitude	Magnitude	Matched	Non-matc	Total	Freq%
0.18574	0.18574	5	0	5	100

Figure 2.5: Example of Excel output for skin conductance data collected from BIOPAC

2.6.5 Heart rate and heart rate variability

Heart rate is defined as the number of beats calculated per unit of time. The changes of heart rate are related to the activity of sympathetic and parasympathetic (See section 2.5 on page 36 for details on sympathetic and parasympathetic activity). The heart rate tended to increase over a short time course (e.g. stress and task) (Demello 1999, Sammito et al. 2015).

Heart rate variability (HRV) is defined as variations of beat-to-beat in length of time between heartbeats (Billman 2011). HRV allows the assessment of physiological changes over a short or long time course, the durations can range from minimum of 5 minutes to 24 hours (Malik et al. 1996, Chandra et al. 2003). The physiological changes are mediated by the activity occurring in the sympathetic and parasympathetic nervous systems (Piéri et al. 2001). A high HRV normally indicates a good and healthy heart. Healthy people with high HRV have been shown to a better way of handling stress. HRV can be an indicator of how the human body responds to stress exposure. It was suggested that the HRV is lower than normal resting when a person has been exposed to acute stress (i.e. mental or physical activity) (Karim et al. 2011, Dong 2016).

There are several HRV parameters that can be used to measure these activities, and the most common types of HRV parameters used in past research of HRV are (i) time domain i.e. standard of all normal RR intervals (SDNN), and (ii) frequency domain i.e. high-frequency (HF) (Billman 2011, Kim et al. 2018). This method seems to be useful to record short term responses of heart rate variability under laboratory settings (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996, Milicević 2005).

SDNN measures an overall activity in autonomic nervous system (Wood et al. 2002). It represents a combination of parasympathetic and sympathetic activity (Malik et al. 1996). However, these activities also can be quantified independently, such as using frequency domain i.e. HF (Libby et al. 2012). But, the HF could represent as a mix of activation of parasympathetic and sympathetic activity (Berntson et al. 1997, Goedhart et al. 2008, van Ravenswaaij-Arts 1993). However, in most cases, the HF used as an indicator of parasympathetic activity (Libby et al. 2012). As reviewed by Castaldo et al. (2015), Rodrigues et al. (2018), exposure to stress could increase heart rate and reduce HRV. It was further suggested that stress also influenced a reduction in HF, and either increased or decreased SDNN (Castaldo et al. 2015).

Data recording:

The heart rate can be recorded using photoplethysmography (PPG) or electrocardiogram (ECG). When recording using the PPG the sensor measures the blood volume using optical method by bouncing infra-red light against the surface of the skin to measure the amount of changes in light that reflected back. The pulses can be detected according to how much blood is in the skin, higher pulses resulted from more blood. Commonly, the pulse sensor was attached on the participants' thumb. Then, when recording using ECG, three pre-gelled disposable electrodes were used to obtain the signal. Two electrodes were placed on left (black cable) and right clavicle (white cable) and just above the left waist (red cable), see Figure 2.6.

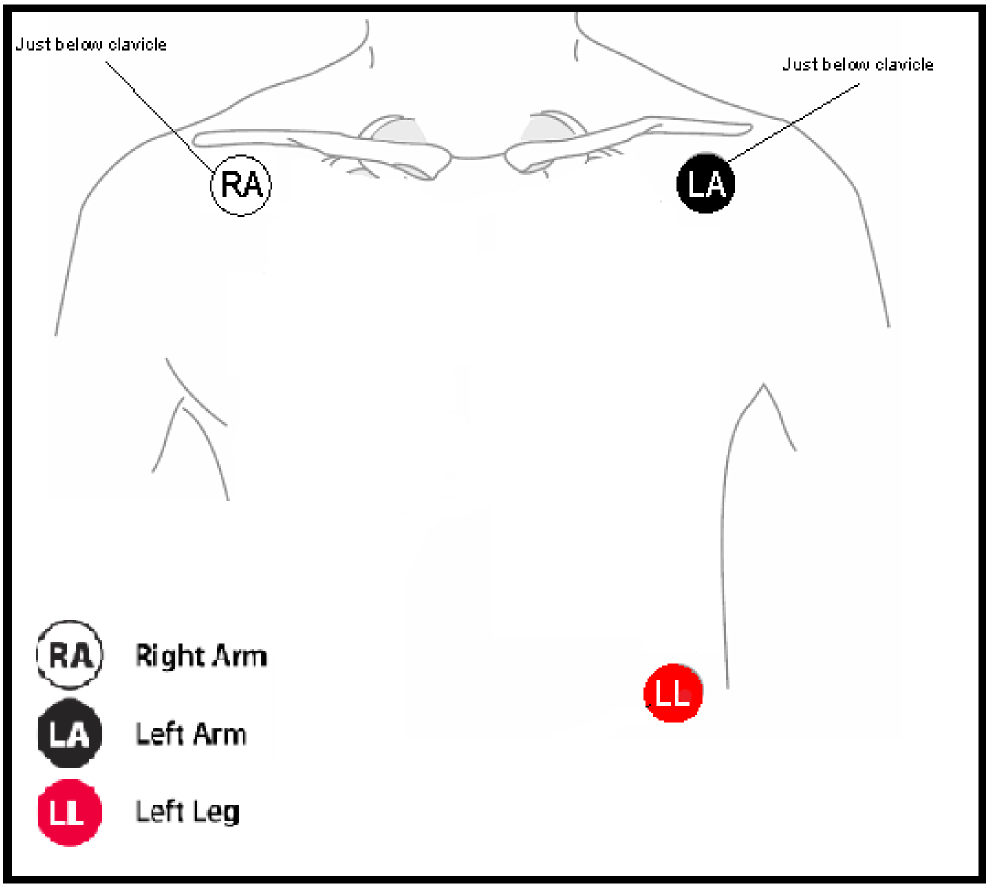


Figure 2.6: Electrocardiogram (ECG) sensors placement (from Cable & Sensor (2010))

Data analysis:

In accordance to recommendation by Tarvainen et al. (2014), the heart rate scores can be pre-processed using Kubios 2.0 software. Kubios is a free software that downloaded from `https://www.kubios.com/`. The signals were corrected using the default option "medium artefact correction" in Kubios software. Additionally, to remove disturbing low frequency baseline trend components, a Lambda value of 500 was set to the "smoothness prior method" option (Luque-Casado et al. 2013). Then the data were manually inspected and corrected for any artefact on or removal (Krypotos et al. 2011).

Then the data from heart rate variability can be pre-processed using AcqKnowledge 4.4 software. The data noises from PPG signal (i.e. movement) were removed using high-pass filters with cut-off frequencies of 0.5 Hz (Lu et al. 2008, Jeyhani et al. 2015).

2.7 Statistical test of data analysis

This section focuses on explaining the selection and use of a statistical test to test the experimental hypotheses.

When investigating quantitative data, data analysis involves the use of statistical measures such as mean (M) and standard deviation (SD) (Watkins 2016). M is a total of scores of a group (x) then divided by the total number of scores of the group (N). The formula is $\frac{\sum x}{N}$. SD is a square root of the sample variance. The formula is $SD = \sqrt{\frac{\sum |x-M|^2}{N}}$. The variance is an average of the square of the differences of the mean. The formula is $Variance = \frac{\sum |x-M|^2}{N}$.

2.7.1 Data cleaning

Before choosing statistical tests to test a hypothesis, data cleaning needs to be done. The purpose of data cleaning is to remove any outliers in the data (unreliable data) that can affect the results (Greer & Mulhern 2001). An outlier is defined as a number that lies far from the overall data distribution. Outlier data should be removed from most or all of the remaining measurements in the data distribution (Shafer & Zhang 2012).

In order to remove or check any outliers in the data, the distributional properties of data have to be checked or observed. The distribution of data in the population can also help to determine statistical analysis that should be used, i.e. either parametric or non-parametric test (Greer & Mulhern 2001). The two tests will be discussed further in the next section.

As suggested by Razali et al. (2011), two tests are commonly used to check the distribution of the data: (i) Shapiro-Wilk test and (ii) Kolmogorov-Smirnov test. According to Razali et al. (2011), both tests are used to test the normality of the distribution of the data. The tests depend on a normal distribution hypothesis as follows: Null hypothesis (H0): The data is not normally distributed. Alternative hypothesis (H1): The data is normally distributed. When testing the hypothesis of data distribution, if the p-value is more than 0.05, the null hypothesis is rejected. The Shapiro-Wilk test is said to be suitable for the data samples smaller than 50, and the Kolmogorov-Smirnov test is suitable for data of more than 50 samples (Razali et al. 2011). However, the Shapiro-Wilk test is more stable than the Kolmogorov-Smirnov test when testing the data distribution, regardless of the size of the data set (Ghasemi & Zahediasl 2012).

Normal data distribution is described as a bell curve and the data is assumed to be distributed in a similar way. The distribution of the data also can be observed based on histogram by looking at the shape of data distribution. Another example of data distribution is skewed data distribution. A skewed distribution is similar to normal data distribution but the data is either stretched out to the left or the right. It can be described as (i) positive skewness (right) or (ii) negative skewness (left).

A way to reduce the skewness in data distribution is by observing the distribution of the data using a box-plot. The box-plot extends from the lower quartile to the upper quartile and any points beyond these limits can be considered as outliers (Greer & Mulhern 2001). It was also suggested by Osborne & Overbay (2004) that any data that more than 3 standard deviations should be removed.

2.7.2 Statistical tests

Based on the results from the tests of normality and skewness, as mentioned earlier, there are two types of statistical test that are commonly used to perform the statistical analysis: (i) *parametric* and (ii) *non-parametric*. The parametric test analyses the data based on the group means, and non-parametric tests analyse the data based on group medians (Gunawardena 2011). The non-parametric test is an alternative test to be used when the data collected did not meet assumptions of parametric tests. This usually occurs when a study has a small sample size of participants (less than 20 participants) (Gunawardena 2011).

Examples of *parametric tests* are the t-test (to compare two groups) and analysis of variance (ANOVA) (used to compare more than two groups). The ANOVA has within-design (repeated measures) and between-design (comparing between groups) and mixed-design (repeated measures and comparing between groups) (Shankar & Singh 2014).

Examples of *non-parametric tests* are a Wilcoxon test (instead of an independent t-test), a Mann-Whitney test (an alternative to the paired t-test), a Kruskal-Wallis test (to replace a one-way ANOVA test) and a Friedman test (Mixed-design test).

In physiological research, repeated measures study design is commonly used (Park et al. 2009). In repeated measures study design, the data is collected at certain time-points (e.g. at pre- and post-measurement time-points). Then when the study involves comparing between two or more groups or conditions, a mixed-design ANOVA test (parametric test) or a Friedman test (non-parametric test) can be used. Both tests analyse if there are any changes or differences in the data across time-points and between groups or conditions.

2.7.3 Statistical analysis using parametric test

This section presents an example of a statistical test using a parametric test. The parametric test is a mixed-design ANOVA (repeated measures and comparing between groups).

Example of experimental design:

The experiment aims to assess participants' heart rate before and after watching a video. There are three conditions in this study: (i) Condition A, (ii) Condition B, and (iii) Condition C. In condition A, participants watch a funny video. In condition B, participants watch a sad video. In condition C, participants watch no video. Participants are randomly assigned to one of the three conditions. The levels of participants' heart rate are taken before and after watching the video.

Hypothesis:

After watching a funny video, participants' heart rate levels will increase compared to before watching the video. It is also hypothesised that participants' heart rate in condition A will be higher than in other conditions.

Variables:

Variable is defined as a measure used in a study that can be analysed such as the time when the heart rate was measured and the levels of heart rate at each time the rate was measured (Shankar & Singh 2014). Based on this example, the repeated measure in this study is heart rate (before and after video). The levels of heart rate of 'before' and 'after' are considered to be the dependent variable. The *dependent variable* is defined as the variable that the researcher is interested to measure in a study (i.e. levels of heart rate). The three conditions in this study are considered as a between-subject factor. A *between-subject factor* is defined as the independent variable. The *independent variable* is defined as a variable that can affect the dependent variable. For example, participants' heart rate may increase after watching a funny video compared to before watching the video.

Example of multivariate analysis:

This section presents an example of results of a statistical analysis test using ANOVA. The analysis can be done using statistical software such as Statistical Package for the Social Sciences (SPSS). If the data is normally distributed, ANOVA is a reliable test to determine the significance of any difference in mean scores.

Assumption of sphericity:

When conducting an ANOVA test, it is important to examine the assumption of sphericity. Sphericity is when the variances of differences between all possible pairs of related conditions (within-subject factor) are equal (Park et al. 2009, Brace et al. 2016). If the assumption of sphericity is not met, a Greenhouse-Geisser correction should be used to make an adjustment to the degrees of freedom in the ANOVA. The degrees of freedom (df) is a value that is related to the number of participants in the whole experiment and the number of values in the sample size that are free to vary (Pandey & Bright 2008, Healey 2014, Brace et al. 2016).

Result:

Example of results for the ANOVA: $F(1,200) = 1.234, p = 0.050, r = 0.900$. The formulae: F(df, df of error(time)) is equal to F-value, p-value, effect size. Definition of the terms are as follows:

- *F-statistic* (F) is defined as the ratio of the two mean squares of between and within groups or conditions (Greer & Mulhern 2001), where they are also described as the variance. The variance is as the result of the manipulation of the between-subject factor. The variance is the difference between conditions or groups, divided by the variance due to error (within-subject factor). The formula is $F = \dfrac{Mean_{between}}{Mean_{within}}$.
- *P-value* (p) is a conditional probability. In a statistical test, the p-value needs to be equal

or less than 0.05 to reach the accepted level of statistical significance (Rothman 1990, Greer & Mulhern 2001).

- *Effect size* (r) is defined as the difference between the mean of two different intervention groups or conditions, divided by the mean of the standard deviation of both groups or conditions (Coe 2002, Sullivan & Feinn 2012). The formula is $r = \frac{x_1 - x_2}{mean\,SD}$. According to Cohen's term (Coe 2002), there are three categories of effect size: (i) small (r-value below or equal to 0.2), (ii) medium (r-value below or equal to 0.5), (iii) large (r-value larger than 0.8).

- *Alpha value* (α): If the p-value is significant, this shows that there was a meaningful difference in mean scores between groups or conditions (Greer & Mulhern 2001). In order to determine the significance level of the p-value, an alpha level (α) needs to be set to 0.05 or 5%. The α level of 0.05 is the common standard used in research. The alpha-level is the probability of the chance of rejecting a null hypothesis.

Multiple comparisons:

However, the p-value does not give precise information about the differences across conditions or time-points, therefore, a further comparison test needs to be done. The further comparison test was done in order to reduce the risk of false-positive results, where the null hypothesis is rejected but the result is not true. One of most used tests in scientific research is the Bonferroni test (Brace et al. 2016). The Bonferroni test performs a multiple comparison test to the p-value obtained earlier. For example, based on the study design described earlier, there are three number of comparisons: Condition A compared to Condition B, Condition A compared to Condition C, and Condition B compared to Condition C. Therefore, the significant p-value will be divided by 3.

Figure 2.7 shows a summary of a sequence for performing statistical data analysis.

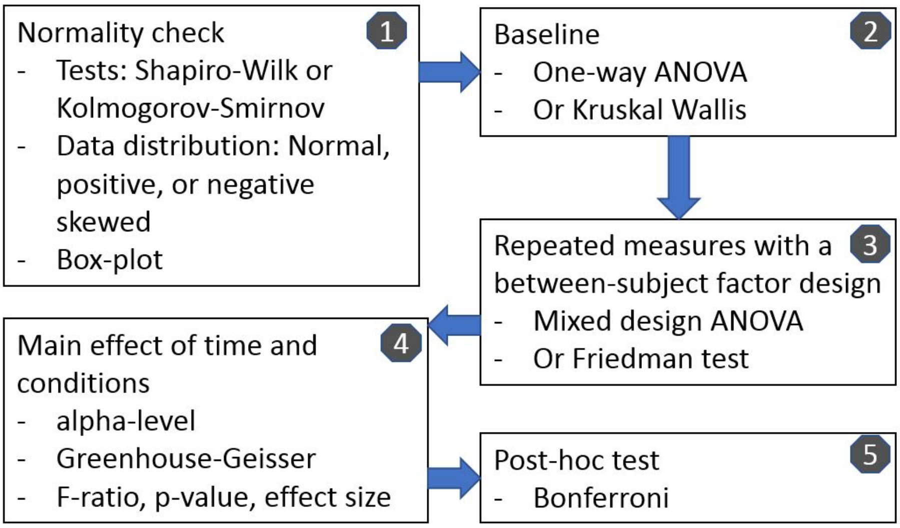


Figure 2.7: A flowchart for commonly used statistical tests.

2.8 Rationale of this thesis

Overall, the previous sections review the literature on animal therapy, pet-robot therapy, Paro robot, stress and methods use to assess human responses, and methods to analyse the physiological responses and understanding the ways of performing the data analysis. Based on review in section 2.3 on page 16, the potential benefits of intervention of Paro have been categorized into three, as follows:

i. Paro as a companion and mediator in *social interaction*.

ii. Intervention of Paro to *reduce pain, anxiety and decrease loneliness*.

iii. Intervention of Paro to *improve psychophysiological health*.

Thus, this section highlights a rationale of this thesis based on a gap and limitations found in Paro literature and some things should be considered for further investigation.

Study sample population

Firstly, previous studies of Paro mostly focused on people that have been diagnosed with long-term mental health problems such as depression and anxiety in general population, and dementia and autism, e.g. Pipitpukdee et al. (2011), Okita (2013), Robinson et al. (2013c). However, there are some limitations to the research that can be undertaken with vulnerable elderly people as mentioned by Robinson et al. (2013b). For instance, participants might have a difficulty in committing to the experiment due to poor physical and mental health, and difficulty with communication. These limitations might lead to an inability for participants to complete the entire experiment.

Healthy young adults

Furthermore, since the Paro robot shows a potential for improving human health, the use of the robots should not be limited to such populations. For example, according to Mowbray et al. (2006), Eisenberg et al. (2007), Hunt & Eisenberg (2010), Jarolmen & Patel (2018), there are numbers of students that seek therapy especially during the examination period. As the examination period can be stressful for students (Trammell 2017). Therefore, the effectiveness of the therapeutic robots in such populations i.e. healthy young adults should be further investigated. Furthermore, an advantage of recruiting students, is that the students are easy to recruit.

At the same time, when investigating the effectiveness of Paro on healthy young adults, there are only a few studies that have looked at this population (Mitsui et al. 2001*a*,*b*, Kawaguchi et al. 2012, Wood et al. 2015). The limitations of these studies are the use of small numbers of participants (except the study by Wood et al. (2015), a lack of comparison between conditions, the use of within-subject experimental design (Mitsui et al. 2001*a*,*b*, Kawaguchi et al. 2012), and a lack of gender-balance (e.g. the majority of the participants were male) in Mitsui et al. (2001*a*). Hence no strong conclusion on the effects of Paro can be made from these studies (Mitsui et al. 2001*a*,*b*, Kawaguchi et al. 2012). The study by Wood et al. (2015) used large number of participants (more than 100 participants). However, the study focused on the effects of Paro as a mediator in dyadic interaction, and not on the use of Paro in the context of reducing effects of acute stress or in understanding how Paro intervention affects physiological changes over time.

Effects of Paro on different sample population

In addition, in terms of interacting with the Paro robot, Robinson et al. (2013c) mentioned that the younger population acted differently when interacting with Paro compared to the older population. Healthy young adults felt more excited when playing with Paro (Mitsui et al. 2001*a*), compared to elderly people with mild or moderate dementia (Robinson et al. 2013c). This may suggests an intervention with Paro may have different benefits or different potential in improving health in both population.

Paro intervention and acute stress

Furthermore, in terms of the effects of Paro on reducing effects of acute stress, only two studies have investigated this, that is McGlynn et al. (2016) and Crossman et al. (2018). These studies investigated the effects of short-term sessions with Paro, during stressful cognitive tasks (McGlynn et al. 2016) and after a stress inducing task (Crossman et al. 2018). However, in the study by McGlynn et al. (2016), participants were not allowed to touch Paro, in order to explore on the effects of Paro being present before, during and after a stressful situation. Meanwhile, Crossman et al. (2018) did not find an improvement in stress reduction after a session with Paro. However, there was a positive result in terms of improving mood after interacting with Paro Crossman et al. (2018).

In addition, these studies also used a different sample population, where participants in McGlynn et al. (2016) study were elderly people aged between 65 and 77, and in Crossman et al. (2018) study were healthy children aged between 6 to 9 years. This shows an obvious age gap in terms of sample population when investigating the effectiveness of Paro on reducing stress responses.

Additionally, different stressors were also used in each study i.e. Raven's Progressive Matrices Test (RPM) (McGlynn et al. 2016) and Trier Social Stress Test (TSST) (Crossman et al.

2018), although both stressors are reliable to induce mild stress in a laboratory-setting. However, the TSST mainly focuses on inducing stress related to social situation and RPM focuses on inducing stress based on cognitive performance. Furthermore, based on positive results found in (McGlynn et al. 2016), it is plausible that interacting with Paro may help in reducing different types of induced stress such as environmental stressor (i.e. noise).

Induced Stress

In the review of noise as stressor in section 2.4 on page 34, it was acknowledge that the exposure to noise can be harmful to mental and physical health. In reviewed by Münzel et al. (2014a, 2018b), there are a few studies claimed that there is a relationship between exposure to noise and cardiovascular diseases. The noise also can potential bring harm to human hearing, sleep, task performance, communication, and emotional states (Westman & Walters 1981). However, it was claimed by Parsons (2002) that most of people not aware that noise also can contribute to one of their daily stressors. Therefore, it is important to understand how the Paro intervention may helps in reducing or minimizing the impact of noise exposures on psychological and physiological functions. This study not only helps to expand the Paro literature (as no previous study of Paro that looking at the effects of Paro intervention on noise exposure), but it will also expand the noise literature and suggest a potential stress reducing intervention for environmental stress.

Effects of Paro on physiological stress responses

Therefore, based on the studies reviewed earlier, no previous study of Paro in a laboratory setting has examined the reducing effects of Paro on induced stress and on physiological responses, in a sample population of healthy young adults. Only one study, Crossman et al. (2018) has looked at the effects of Paro on reducing the effects of induced stress, in the context of physiological responses and laboratory settings. However, Crossman et al. (2018) only measured the reactivity in hypothalamic pituitary adrenal (HPA) on stress responses using salivary cortisol, and using a population sample of children.

Despite that, based on the literature, there are a few studies that have found an improvement in physiological responses after interacting with Paro (Mitsui et al. 2001*a*,*b*, Wada & Shibata 2007b, Kawaguchi et al. 2012, Robinson et al. 2013c, Carretero & Sáez 2017, Petersen et al. 2017, Crossman et al. 2018). But, some of these studies used small sizes of participants (n = 6 to 60). According to Bethel & Murphy (2010), the largest sample of participants for human study should be more than 60. The largest sample size can help to give an adequate power of effects to detect differences when doing a group comparisons.

Extraneous variables

Furthermore, most of the previous studies of Paro were done in natural settings such as nursing homes, and the interactions with Paro were in groups of people or with the researcher or the caregiver present (Robinson et al. 2013b). The presence of other people and other objects during the session can be identified as extraneous variables and may have had an effect on the social activity. For example, in studies by Kidd et al. (2006), Wada & Shibata (2007a), participants were given a chance to play with Paro freely in a public area or open space in nursing homes. There might be other objects in the area that also influenced any effects that occurred during the interaction with Paro. Therefore, a sole effect of Paro and the effect of interaction with it could not be fully understood.

Another extraneous influence is the humidity of in the public settings. As mentioned by Boucsein (1992), the temperature of the room could influence the physiological responses. Therefore, when studying in natural settings, it can be difficult to measure or control the temperature compared to doing an experiment in a lab. A previous study by Palestrini et al. (2017) mentioned that participants' heart rates increased due to high temperature in the ward, hence reducing the effects of intervention therapy on physiological stress responses. These extraneous variables need to be carefully controlled by investigating the stress reducing effects of Paro on induced stress in more controlled environment such as laboratory settings, where the temperature is unlikely to vary.

In conclusion, based on the summary of literature, the uses of therapeutic robots have been shown to have potential for improving people's health, in both mental and physical health. However, there are some limitations to the research and some unresolved methodological issues. Furthermore, if therapeutic robots are shown to reduce stress in vulnerable population, then their use with healthy young adults who do not have dementia but who suffer from stress should be investigated. The thesis aims to add evidence to the existing literature, as no study of the stress reducing effects of Paro on induced stress in laboratory settings, in healthy young adults, and in physiological responses, has been reported before.

Chapter 3

Study 1: Comparing effects of interacting with an active and inactive Paro

3.1 Overview of Study 1

As reviewed in Chapter 2, pet robot therapy or therapeutic robots have shown a lot of potential for improving mental health in diverse groups such as the elderly and children (Okita 2013, Wada & Shibata 2007b). Emerging technologies such as the Paro have been used commercially in care homes. Hundreds of Paro have been sold internationally for therapeutic purposes (Levsen 2015). Although there can be remarkable positive psychophysiological improvement as the results of interacting with Paro robot, most of studies using young healthy adults (ages between 18 to 35 years (Petry 2002, Dyrbye et al. 2006)) were primarily exploratory (Mitsui et al. 2001*a*,*b*, Kawaguchi et al. 2012). These studies also used a very small sample size (less than 10 participants) and lacked controlled condition comparisons (within-subject study design).

According to World Health Organization (2014), the adult population is often subject to mental health problems such as stress, anxiety, depression that can increase the risk of cardiovascular diseases. As reviewed in section 2.4, exposure to acute stress in daily life can induce negative moods and expose people to the risk of cardiovascular diseases. Improvements in psychophysiology are important for reducing stress, and short-term improvements could lead to better long-term mental and physical health. Therefore, these reasons motivate Study 1 to investigate further about the potential of pet robot therapy such as Paro in reducing stress on healthy adults or non-clinical population.

3.2 Main aims

The aim of Study 1 was to investigate the effects of short-term interactions with Paro under laboratory settings in healthy adults.

A further aim was to understand whether interacting with a Paro could help to reduce the level of stress response to previously encountered noise.

Additionally, another aim was to examine which characteristics of the Paro and humans might be responsible for any effects during interventions, or for reducing the temporarily induced stress. The primary experimental manipulations were the physical embodiment and behaviour of Paro (e.g. appearance, soft fur, animated movement, voice), and human behaviours towards the Paro (i.e. talking, stroking, and petting).

The effects were measured using observed facial expressions, self-reported stress, positive and negative affect, and physiological response i.e. skin conductance.

In order to investigate the effects of the Paro to reduce stress in laboratory settings, a laboratory stress-inducing event was used to assess the psychophysiological responses. In Study 1, a sound consisting of a combination of noises was used to induce mild stress. Study 1 was designed in controlled settings to focus on the immediate effects of interacting or presence of Paro on reducing acute stress such as noise.

3.3 Terms

The terms listed below are defined for the scope of Study 1:

- *Time-point* refers to the time when the physiological measure was taken. There are four time-points in this study where time-point 1 labelled as t1, time-point 2 labelled as t2, time-point 3 labelled as t3 and time-point 4 labelled as t4.
- *After time-point* refers to the time when the participants need to fill in the questionnaires. Participants fill in the questionnaires after each physiological measures time-point. There are four after time-points in this study where after time-point 1 labelled as at1, after time-point 2 labelled as at2, after time-point 3 labelled as at3 and after time-point 4 labelled as at4.
- *Session 1: noise* refers to a stress session at time-point 2, where the participants were asked to listen a noise.
- *Session 2: Paro* refers to intervention session with a robot at time-point 3. There are four conditions in this study where Condition 1 labelled as C1, Condition 2 labelled as C2, Condition 3 labelled as C3, and Condition 4 labelled as C4.
- *Changes in scores* refers to new variables created for changes in scores between time-points. The variables for self-reported responses labelled as at2mat1 (changes in scores between session 1: noise and pre-measurement), at3mat1 (changes in scores between session 2: Paro and pre-measurement), and at3mat2 (changes in scores between session 2: Paro and session 1: noise). The variables for physiological responses labelled as t2mt1 (changes in scores between session 1: noise and pre-measurement), t3mt2 (changes in scores between session 2: Paro and session 1: noise), 4mt3 (changes in scores between post-measurement and session 2: Paro), and t4mt1 (changes in scores between post-measurement and pre-measurement).

3.4 Summary of Study 1

This section provides a summary of methods used in Study 1. In the next section 3.5 provides more details about important consideration of this study.

Methods:

Study 1 used a repeated measures and between-subject design. Study 1 had four time-points (t1 - 4) and four conditions (C1 - 4). The physiological measure (i.e. skin conductance) was collected at four time-points: A pre-measurement (t1) was followed by Session 1: a 3 minutes stress-inducing noise (t2), Session 2: a 3 minutes intervention session with the Paro robot (t3) and a final measurement (t4).

Then self-reported stress, and positive and negative affect were collected after pre-measurement (at1), after session 1: noise (at2) and after session 2: Paro (at3).

Then in order to observe facial expressions towards the robot, participants' session with the robot (during session 2 at t3) were recorded on video.

Conditions: Participants were randomly assigned to one of four conditions (C1 - 4): Condition 1 (C1): a full interaction (i.e. talking and stroking) with an active Paro, Condition 2 (C2): talking to an active Paro, Condition 3 (C3): stroking an inactive Paro, and Condition 4 (C4): no interaction with an inactive Paro).

A detailed experimental procedure for Study 1 is provided in section 3.9 on page 75.

3.5 Important considerations and rationale of Study 1

The following are the important considerations that Study 1 aimed to explore.

3.5.1 Short-term interaction

A 3 minutes intervention session duration was selected because past research of animal or Paro therapy has found that a few minutes session with a pet i.e. 2 to 5 minutes can result in an improvement in moods (Shiloh et al. 2003) and physiological responses (Mitsui et al. 2001*b*), and stimulate brain activity (Kawaguchi et al. 2012), and result in positive observed facial expression (Takayanagi et al. 2014).

In Study 1, the physiological response was measured using skin conductance response (SCR), as an indicator to autonomous nervous system activity (Figner & Murphy 2011).

Therefore, a *hypothesis 1 (H1) hypothesised that during a short-term with full interaction with an active Paro (C1) compared to other three conditions: talking to active Paro (C2), stroking an inactive Paro (C3) and no interaction (C4) would result in greater effects on H1(a) improving moods and H1(c) positive facial expression.*

3.5.2 Key factors

There are a number of key factors that could lead to improvement in moods and physiological responses during short-term interaction with pet robot therapy. The factors are *(i) features of Paro* such as its appearance, physical embodiment, and interactiveness, and (ii) *human behaviour* towards the robot (i.e. talking, stroking, smiling and eye gaze).

(i) Robot features:

The Paro has several features such as a cute face, soft fur, body movement, cute voice and physical embodiment. These features could influence moods and encourage engagement during interacting with the robot, that lead to the positive improvement mentioned in past research of Paro (Barker et al. 2003, Sherman et al. 2009, Kolling et al. 2016) (See section 2.3, page 16). The positive moods are suggested to be related to a reduction of the induced stress (Crossman et al. 2018). It has also been suggested that the momentary enhancement in positive moods could benefit mental and physical health (Gallego-Perez et al. 2014). However, each feature of Paro could have a different influence on moods and physiological responses (Mitsui et al. 2001*b*, Takayanagi et al. 2014, Crossman et al. 2018). Therefore, to investigate the different effects of Paro's features, an inactive Paro (in C3 and C4) was used to compare with the effects of interacting with an active Paro (C1 and C2).

(ii) Human behaviour:

Another aspect that could influence the moods and physiological stress responses is human behaviour towards the robot such as stroking, talking and eye gaze. These behaviours were suggested to be linked to user engagement with pet robot in social support and helping to buffer or reduce the stress response (Horsten et al. 1999, Ditzen et al. 2007). Ditzen et al. (2007) mentioned that stroking or physical contact will reduce blood pressure and the effects that are mediated by autonomic nervous system (ANS) and hypothalamic pituitary adrenal (HPA). In addition, in terms of therapy, talking to someone can also help in reducing the physiological responses to stress (Horsten et al. 1999). However, in a review by Beck & Meyers (1996) on animal-therapy, the authors mentioned that a lower level of stress was found in people who talked with pet compared to human. Based on these benefits, Study 1 compared the effects of different human behaviour towards Paro on moods and physiological responses.

Two conditions were included in order to examine effects of talking to an active Paro (C2) and stroking an inactive Paro (C3) might have on moods and stress responses compared to full interaction with active Paro (C1).

There is also a possibility that the participants would feel better at the end of experiment without any interaction with Paro. Therefore, condition 4 will be used as baseline or control condition. Participants will be asked to sit back and relax. Paro will simply be present on the table. Participants are not allowed to touch or talk with the Paro, and the Paro will be switched off. If a similar improvement in stress or mood is found in C4 compared to the other conditions, this could suggest that the presence of Paro in the room could influence the improvement. This is because the 'cuteness' of Paro can be represented as a calming visual object (Sherman et al. 2009, Kolling et al. 2016). It was suggested that by just watching a calm visual object such as a plant (Dijkstra et al. 2008), fish or aquatic plants in aquarium (Barker et al. 2003, Buttelmann & Römpke 2014) or watching a video of cute pet (Myrick 2015, Kogan et al. 2018) may have positive effects and reduce stress. Another possibility is just resting might reduce the participants' stress response.

3.5.3 Skin conductance responses

Study 1 measured the changes in physiological response using skin conductance response (SCR). *It was hypothesised that during a full interaction with an active Paro (C1) at session 2 versus the other conditions: talking to active Paro (C2), stroking an inactive Paro (C3) and no interaction (C4) would result in: Hypothesis 1, H1(b) higher skin conductance responses (at t3).*

The prediction was based on findings from a previous study of Paro by Mitsui et al. (2001*a*). As discussed in literature review chapter on page 25, Mitsui et al. (2001*a*) found a significant increase in skin conductance responses when stroking a column covered with Paro's soft fur and vinyl, and when interacting with an active and an inactive Paro. The SCR were significantly higher in active Paro compared to inactive Paro condition, and the higher SCR in active Paro condition was associated with positive moods, as reported by participants (Mitsui et al. 2001*a*).

Then in conjunction to first study by Mitsui et al. (2001*a*), Mitsui et al. (2001*b*) did a comparison between hand movement, stress task, stroking an active and an inactive Paro. They found a significant increase in SCR from baseline to stress task session (mental math), there was also an increase trend in SCR when stroking active and inactive Paro. Whilst the study by Mitsui et al. (2001*b*) tried to distinguish between positive and negative arousal in SCR based on attitudes towards Paro, no self-reported response were measured in the study. The author also did not report a statistical group comparison between the negative and positive arousal. The study by Mitsui et al. (2001*b*) also had some limitations such as small number of participants, and used within-subjects study design. Although some of the results showed statistical changes in scores, a larger magnitude of effect was not detected between conditions. The results may also have been influenced by personal bias or fatigue during the experiment (Charness et al. 2012).

Furthermore, no previous studies of Paro have looked at differences in physiological arousal to Paro, to an environmental stressor (i.e. noise) or whether interacting with Paro reduces the effect of noise induced stress.

Therefore, Study 1 tested that, based on changes in scores between baseline and final-measurement. Hypothesis 2 (H2): A full interaction with Paro compared to other three conditions following exposure to a stress-inducing session would result in H2(a) an improvement in mood and H2(b) a reduction in skin conductance responses.

This prediction was based on findings in animal therapy that investigate stress-reducing effects of interacting with animal on skin conductance responses (Allen et al. 1991, O'haire et al. 2015). The skin conductance response (SCR) tended to reduce or to return to baseline following stress exposure, as a result of recovery (Ulrich et al. 1991).

Research examining the physiological effects i.e. SCR of environmental stress have shown that noise has the ability to increase SCR (Alvarsson et al. 2010, Mackersie & Cones 2011, Capobianco et al. 2018). But, based on the literature review on Paro, it was noted that a mental task showed a greater increase in SCR than stroking an active or an inactive Paro (Mitsui et al. 2001*b*). This leads to other questions about whether other stress inducement such as noise will result in similar or higher levels of physiological arousal compared to Paro, and whether interacting with the Paro reduces the physiological arousal resulting from the stress.

3.6 Hypotheses

Based on important considerations discussed in earlier section, the following are the hypotheses of Study 1:

Hypothesis 1 (H1):

The hypothesis 1 (H1) was tested based on scores at session 2: Paro (at t3 or at3).

The H1 was divided into there parts: H1(a) tested based on effects of condition on self-reports, H1(b) tested based on effects of condition on skin conductance responses, and H1(c) tested based on effects of condition on observed facial expressions.

The hypothesis aims to answer whether different aspects of Paro or different behaviour towards Paro may influence participants' responses during intervention session.

It was hypothesised that a full interaction with an active Paro (C1) versus the other conditions: talking to active Paro (C2), stroking an inactive Paro (C3) and no interaction (C4) would result in:

H1(a) higher self-reported positive affect, and lower self-reported stress and negative affect (at at3).

H1(b) higher skin conductance responses (at t3).

H1(c) higher frequency of observed positive and lower negative facial expressions (at t3).

Hypothesis 2 (H2):

The hypothesis 2 (H2) aims to answer whether different aspects of Paro or different behaviour towards Paro may influenced participants' responses, based on changes in scores between baseline and final-measurement. (at3mat1 or t4mt1: In order to assess if the changes of scores between paired of time points differ across four different conditions, new variables were created. The new variables are mentioned in section on page 81, 84, 86, and 93).

The H2 was divided into two parts: H2(a) tested based on effects of condition on self-reports, and H2(b) tested based on effects of condition on skin conductance responses.

It was hypothesised that a full interaction with Paro compared to the other three conditions following exposure to a stress-inducing session would result in:

H2(a) an increase in self-reported positive affect and a reduction in self-reported stress and negative mood (at3mat1).

H2(b) a reduction in skin conductance responses (t4mt1).

3.7 Methods

3.7.1 Participants

72 participants were recruited in Study 1 (37 females, 35 males, participants were aged between 18 to 44, $M = 24.06, SD = 5.384$). Table 3.1 shows summary of age and gender for each condition. A one-way between-subjects ANOVA test (gender as dependent variable and condition as a between-subjects factor) found there was no significant difference in the distribution of gender between conditions, $F(3,68) = 0.53, p = 0.984, r = 0.002$. A one-way between-subjects ANOVA test also (age as dependent variable and condition as a between-subjects factor) found there was no significant difference in age distribution between conditions, $F(3,68) = 1.227, p = 0.307, r = 0.051$.

Table 3.1: Report of age and gender by condition for Study 1

Condition	N	Age	Gender
1	18	M = 24.94, SD = 5.185	M = 8, F = 10
2	18	M = 22.39, SD = 3.837	M = 9, F = 9
3	18	M = 25.44, SD = 5.512	M = 9, F = 9
4	18	M = 23.44, SD = 6.546	M = 9, F = 9
Total	72	M = 24.06, SD = 5.384	M = 35, F = 37

Ethical approval was obtained from University of Sheffield, Department of Computer Science Research Ethics Review Committees (Reference number: 006480). Experiments were conducted in a listening room at Department of Computer Science, University of Sheffield (See Figure 3.1 on page 73 for an overview of the experiment room).

Participants were recruited through University of Sheffield volunteering distribution-emailing list. A small announcement was also placed on Student Services information Desk (SSID). Prior to the experiment, participants were given an explanation about the study (Appendix A.4: Information Sheet). A signed consent form was obtained from participants. Participation was voluntary and all information was treated as confidential. Participants were told that they were free to leave at any time during experiment. Participants were offered a chance to win a £50 gift voucher.

Exclusion criteria

In order to ensure participants' safety when performing the experiments and to limit any factors that will confound the results some exclusion criteria were used. When advertising the study, participants who had (i) skin problems, (ii) heart problems and high blood pressure, (iii) mental health problems, (iv) hearing difficulties, and (v) taking prescribed drugs medication were advised not to participate.

There were several reasons for the exclusion criteria: In this study, skin conductance measures were used to collect physiological reactions. The procedure required participants to sit comfortably and participants were required to have skin conductance sensors attached directly to their fingertips in a material similar to a foam-like plaster. Allergic reactions to adhesive tape and skin problems might occur in persons with sensitive skin. A small amount of electricity was applied through the conductance, therefore, people with pacemakers or heart problems were not recruited in this study. In addition, as this study is related to stress assessment, persons who have high blood pressure or heart problems were also excluded. Next, those who had been diagnosed with hearing problem either one sided or both sided hearing loss also were excluded. In this study, participants need to be able to hear the auditory stimulus at the required level. This is also to avoid the noise increasing the risk of damage to their hearing. Participants taking medicine such vasoactive drugs to increase their heart rate and blood pressure were also excluded.

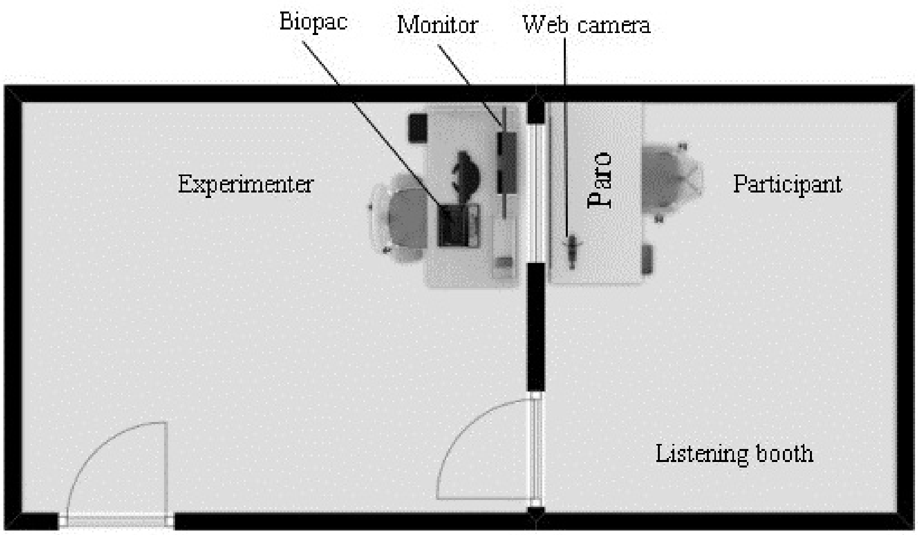


Figure 3.1: Listening room at Department of Computer Science. Notes: This is an overview of the listening room and no exact measurement of the room was included. Participants were asked to sit inside the listening room, and the experiment was monitored by the experimenter from outside of the room. The measuring equipment (BIOPAC) used to collect the physiological data and a monitor (that the participants used to read the experimental instructions and questionnaires were displayed on a 17-inch computer monitor), were on the table at the experimenter side. At the participant side, there was a keyboard, a mouse and a web camera on the table, and the Paro was put on the table during the session 2: Paro.

3.8 Measurement and materials

3.8.1 Pet-therapy robot, Paro

A detailed description of Paro has been provided earlier in this thesis, see section 2.3.

3.8.2 Noises as a stressor

The noises used in the experiment were downloaded from YouTube (www.youtube.com/watch?v=E_KPJhd857w, and www.youtube.com/watch?v=mIXpLA4Q1Tc). The decibel levels of noise were edited using Mp3Gain software (http://mp3gain.sourceforge.net/). Then, to produce different levels of decibel in 3 minutes duration for each sound, the Audacity software were used (http://audacityteam.org/). The approximate levels of decibel were measured using the CR:160 Series Optimus Red sound level meter. The noises were delivered at an approximate volume of 60 to 80 decibels through Sennheiser HD 202 headphones that connected to a laptop.

3.8.3 Positive and Negative affect questionnaire

Positive and Negative Affect Schedule (PANAS) was used to measure the positive affect (PA) and negative affect (NA) of participants in order to confirm their responses to noise and conditions assigned in this study. The questions asked participants to rate their current feelings. There are 20 questions included in PANAS (Refer to Appendix A.8: PANAS). Participants were asked to rate their feelings and moods on a five point Likert Scale (1 to 5) from *Very slightly or not at all = 1, A little = 2, Moderately = 3, Quite a bit = 4, and Extremely = 5*. The scores can range from 10 to 50. The higher scores equate to stronger feelings and moods of positive or negative affect (Crawford & Henry 2004). The PANAS can be used to measure general distress, dysfunction and state anxiety.

3.8.4 Self-reported ratings of stress

Self-reported ratings were used as a subjective measure of levels of stress. Participants were asked “*How do you feel right now?*” at each time-point (after pre-measurement (at1), after session 1: noise (at2), after session 2: Paro (at3). Participants rated their feelings using 7-point Likert-scale ranging from “Not stressed at all = 1” to “Extremely stressed = 7”.

3.8.5 Physiological measurement

Physiological measurement: skin conductance was recorded using the BIOPAC MP36. In this study, the skin conductance sensors had sampling rate of 200 Hz, as this study collected a 3 minutes of sample for each time-point. Skin conductance was recorded using finger sensors that were attached directly to participants' fingertips in a material similar to a foam-like plaster. Figure 3.2 shows the BIOPAC hardware, skin conductance sensors and their placement on the fingers. A detail about skin conductance has been discussed in section 2.6.4 on page 43.

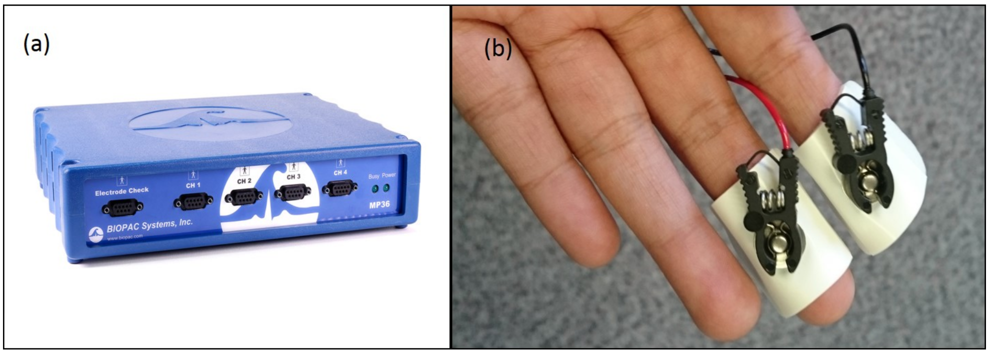


Figure 3.2: (a) BIOPAC MP36 data acquisition unit (BIOPAC 2017), (b) Skin conductance placement

3.8.6 Video recording

Videos were recorded during the 3 minutes that participants were exposed to or interacted with the Paro. The videos were recorded using Sony Z3 mobile phone camera.

3.9 Experimental procedure

Before experiment began, an information sheet (Appendix A.4) was sent through email to participants who confirmed their appointment. Participants were asked to sit in the sound insulated listening room. Inside the room, there was a desk and chair. Then exact instructions were given orally and in printed form to the participants. The explanation depended on the condition assigned to participants. Participants were also asked to follow all the instructions displayed on-screen. After that, participants filled out a consent form and their age and gender. The instructions and questionnaires were displayed on a 17-inch computer monitor. Next, skin conductance reusable electrodes were attached onto participants' finger tips of the index and middle fingers. The electrodes were secured with micro-porous tape. Then participants were left alone in the room.

The experiment consisted of four time-points: pre-measurement (t1), session 1: noise (t2), session 2: Paro (t3) and post-measurement (t4). There are also 4 conditions in this experiment and the conditions differed at the session 2: Paro (t2) only. Participants were randomly allocated to Condition 1 to 4.

In Condition 1 (C1), they were asked to interact with an active (switched on) Paro by stroking and talking to it. In Condition 2 (C2), participants were asked to interact with an active Paro without touching it. In Condition 3 (C3), Paro was inactive (switched off) and participants were asked to stroke it. In Condition 4 (C4), Paro was turned off and participants were asked to sit back and relax. The switched off Paro ensured this condition was the same as the others apart from the lack of any interaction between the Paro and the participants.

The four conditions were labelled as follow: C1:active_stroke_talk, C2:active_talk, C3:inactive_stroke, C4:inactive_rest.

A pre-measurement or baseline physiological responses (t1) was taken for 3 minutes while participants were asked to sit back and relax. After pre-measurement (t1) they were asked to fill out questionnaires: positive and negative affect and self-reported levels of stress (at1). Then they were required to listen to a noise for 3 minutes (t2). Physiological recording was taken during the session 1: noise (t2). After listening to the noise, they required to fill in the same questionnaires again (at2). Then they took part in a session with the Paro for 3 minutes (t3). Physiological recording was also taken during session 2: Paro (t3). After the 3-minute session 2: Paro participants completed the questionnaires again (at3). Their skin conductance response (SCR) was recorded for a further 3 minutes (post-measurement (t4)). The whole experiment was conducted in the following sequence, refer to Figure 3.3. All their interactions with Paro were recorded using a mobile phone camera. Then all the SCR electrodes were removed from participants. Then they were given a chance to ask further questions that related to the experiment. The experiment lasted around 30 to 45 minutes for each participant.

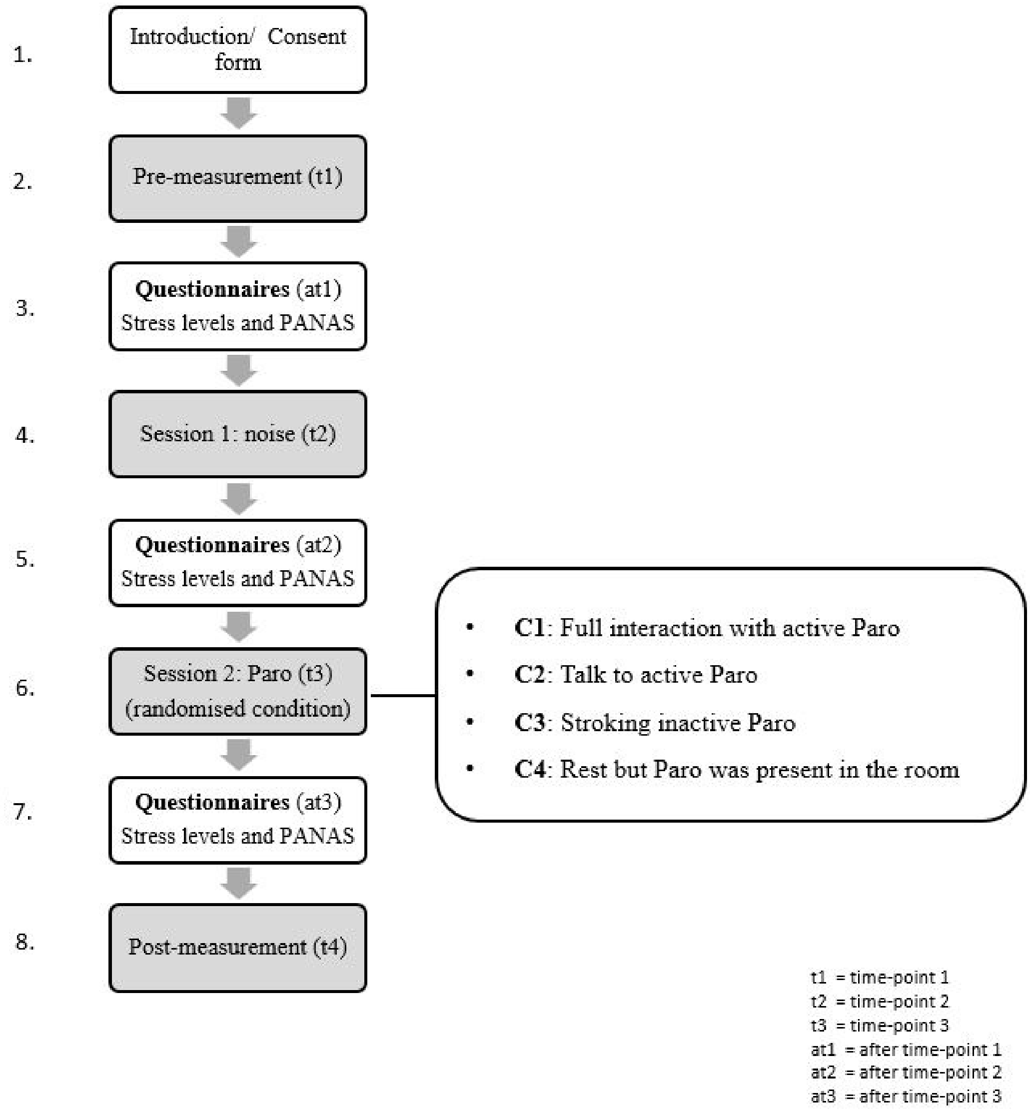


Figure 3.3: Study 1 experimental procedure

3.10 Data analysis

Skin conductance data pre-processing: Skin conductance data were calculated by totalling the number of peaks that occurred during 3 minutes of recordings for each time point (pre-measurement (t1), session 1: noise (t2), session 2: Paro (t3), post-measurement (t4)).

Normality check: Data analysis was done separately for physiological and psychological measures. All data were checked for normality using Shapiro-Wilk test (Razali et al. 2011). The data were carefully inspected using box plots (Osborne & Overbay 2004). Based on the inspection, any extreme outliers were removed.

Baseline: First, the data were analysed in order to check that there was no difference between the participants in the four conditions (C1:active_stroke_talk, C2:active_talk, C3:inactive_stroke, C4:inactive_rest) at the pre-measurement for all variables (skin conductance responses, Positive and Negative Affect Schedule (PANAS), stress levels). A one-way between-subjects ANOVA was performed with pre-measurement (taken at t1 for physiological responses and at at1 for self-reported responses) as dependent variable and conditions as a between-subjects factor. The alpha level was set at $p < 0.05$.

Main effects of time and conditions: Then the data were analysed to look at main effects of time, main effects of conditions, and an interaction effect of time and conditions on the physiological and self-reported responses. A mixed between-within subjects ANOVA was performed for each measure with four conditions (C1:active_stroke_talk, C2:active_talk, C3:inactive_stroke, C4:inactive_rest) as a between-subjects factor and four time points (t1, t2, t3, t4) as within-subjects variable factor. The dependent variables were self-reported stress, positive affect, negative affect and skin conductance. For each analysis, any ANOVAs variables found to violate the assumption of sphericity were corrected with a Greenhouse-Geisser correction and all post-hoc tests were done with a Bonferroni correction. The alpha level was set at $p < 0.05$.

New variables: For further analysis, in order to assess if the changes of scores between pairs of time points differ across four different conditions, new variables were created based on pairs of time points. A one-way between-subjects analysis of variance was performed with each new variable as dependent variable and conditions as a between-subjects factor. A further explanation is reported on each subsection of the results.

Analysis for observed facial emotional expressions:
The video observation was made on 3 minutes of video recording during the session with Paro. There were four facial expressions that were checked and the facial expressions were divided into positive emotion (*smile and laugh*) and negative emotion (*boredom and disgust*). The 3 minutes of videos were divided into 18 blocks of 10 seconds.

All 3 minutes videos were analysed by two evaluators for inter rater reliability. Both evaluators recorded the total number of times each of the four facial expressions occurred during the 3 minutes. The inter rater reliability for all emotions were analysed using Intraclass Correlation Coefficient (ICC) (Landers 2015).

For further analysis, both scores from both the evaluators were averaged, providing a score of the number of times each of the four facial expressions occurred during each videoed session (Takayanagi et al. 2014).

Then a one-way between-subjects ANOVA was performed for each emotion. The dependent variables were smile, laugh, disgust and bored. The conditions as a between-subjects factor. The alpha level was set at $p < 0.05$.

A further analysis was done to compare scores for positive and negative facial expressions for each condition. A new variable, positive facial expressions which totalled up scores of smile and laugh coded facial expressions (Takayanagi et al. 2014). Another new variable is negative facial expression which totalled up scores of disgust and boredom coded facial expression.

Then a one-way between-subjects ANOVA was performed for each new variable. The dependent variables were positive and negative facial expressions. The conditions as a between-subjects factor. The alpha level was set at $p < 0.05$.

Effects of conditions during session 2: Paro In order to test a difference between conditions for each measure, a one-way between-subjects ANOVA was performed with the session 2: Paro (taken at t2 for physiological responses and at at2 for self-reported responses) as dependent variable and conditions as a between-subjects factor. The alpha level was set at $p < 0.05$.

3.11 Study 1: Results

3.11.1 Baseline of each measure

There was no significant difference between conditions at pre-measurement (t1 or at1) in any of the self-reported and physiological responses. These results indicate that participants in all conditions had a similar level of self-reported stress, positive and negative affect and physiological responses at the beginning of the study, all $p > 0.05$.

3.11.2 Self-reported stress

A mixed between-within subjects analysis of variance showed there was a significant effect of time on stress levels, $F(1.693, 115.13) = 78.75, p < 0.0001, r = 0.537$, and no significant effect of conditions, $F(3, 68) = 1.932, p = 0.133, r = 0.079$. There was no significant interaction between condition and time on stress levels, $F(5.080, 115.13) = 2.058, p = 0.075, r = 0.083$.

Pairwise comparisons between time points were done using Bonferroni correction test. As can be seen in Figure 3.4, the stress level scores showed a significant increase after listening to the session 1: noise (at2) from pre-measurement (at1), $p < 0.0001$. The stress level scores then showed significant decrease after the session 2: Paro (at3) from the session 1: noise (at2),

$p < 0.0001$. However, there was no significant difference between after pre-measurement (at1) and session 2: Paro (at3), $p = 1.000$.

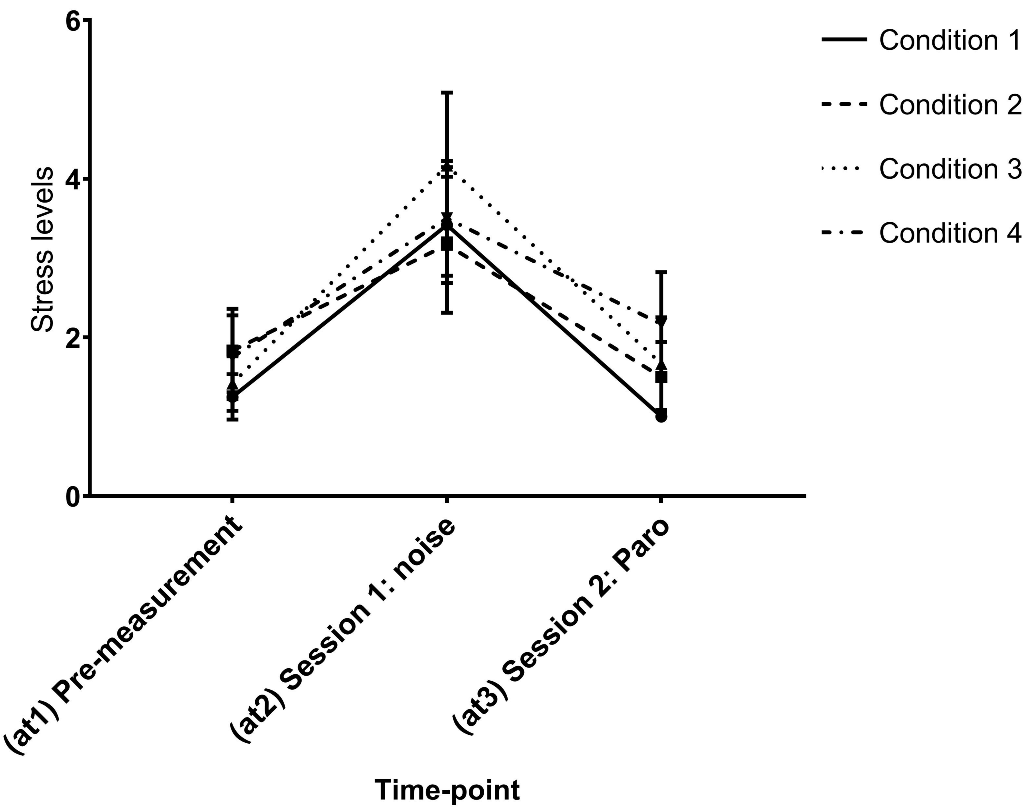


Figure 3.4: The mean with 95% confidence interval for self-reported ratings on levels of stress to each condition.

New variables were created in order to assess differences between conditions in terms of changes in stress level scores. The new variables are $SL\_at2mat1$ = session 1: noise minus pre-measurement, $SL\_at3mat2$ = session 2: Paro minus session 1: noise and $SL\_at3mat1$ = session 2: Paro minus pre-measurement.

There was no significant difference between conditions for $SL\_at2mat1$, $F(3,68) = 1.966, p = 0.127, r = 0.080$. As can be seen in Figure 3.5, it was found that the stress level scores increase from pre-measurement (at1) to the listening to the noise (at2) for all conditions.

There was no significant difference between conditions for $SL\_at3mat2$, $F(3,68) = 2.538, p = 0.064, r = 0.101$. As is apparent in Figure 3.5, it was found that the stress level scores decrease from session 1: noise (at2) to the session 2: Paro (at3) for all conditions.

There was no significant difference between conditions for $SL\_at3mat1$, $F(3,68) = 1.550, p = 0.210, r = 0.064$. No changes in stress level scores were found between pre-measurement (at1) and session 2: Paro (at3) for all conditions, as shown in Figure 3.5.

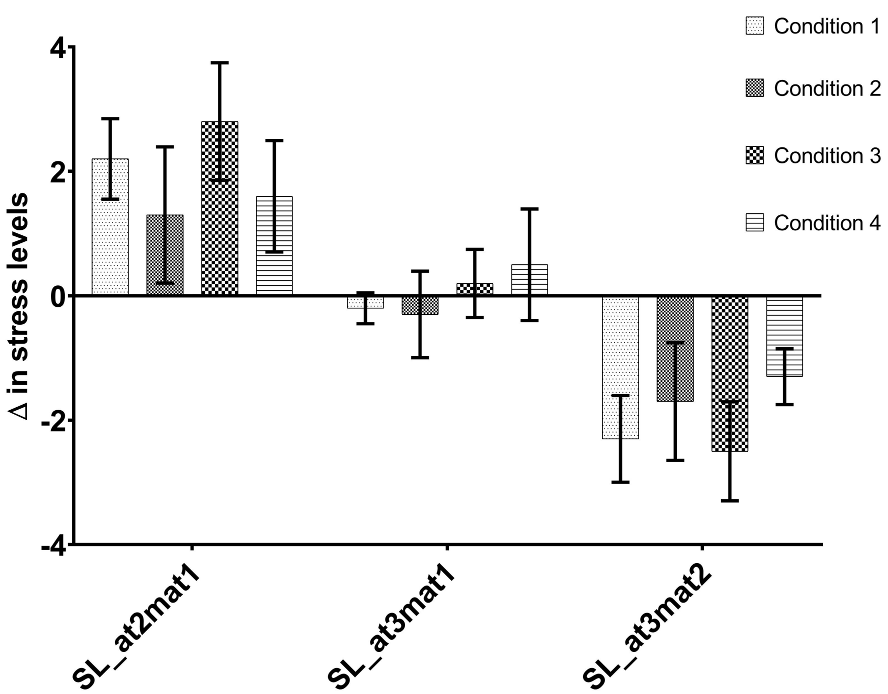


Figure 3.5: The mean with 95% confidence interval for changes of scores on levels of stress to each condition.

Effects of conditions during session 2: Paro

A one-way between-subjects ANOVA showed there were significant difference between conditions at session 2: Paro, $F(3,68) = 4.261, p = 0.008, r = 0.158$. Self-reported stress levels were significantly lower in C1:active_stroke_talk than C4:inactive_rest, $p = 0.004$, as can be seen in

Figure 3.14, page 95. However, differences between the other conditions in self-reported stress levels were not significant.

3.11.3 Positive affect in PANAS

A mixed between-within subjects ANOVA showed there was a significant effect of time on positive affect, $F(2,136) = 10.75, p < 0.0001, r = 0.137$, and no significant effect of conditions, $F(3,68) = 0.641, p = 0.591, r = 0.028$. There was no significant interaction between condition and time on positive affect, $F(6,136) = 1.483, p = 0.189, r = 0.061$.

Pairwise comparisons between time points were done using the Bonferroni correction test. As shown in Figure 3.6, there was a significant decrease in positive affect scores after listening to the noise (at2) compared to the pre-measurement (at1), $p < 0.0001$. The positive affect scores then showed a significant increase after the session 2: Paro (at3) from the session 1: noise (at2), $p < 0.001$. However, there was no significant difference between pre-measurement (at1) and session 2: Paro (at3), $p = 1.000$.

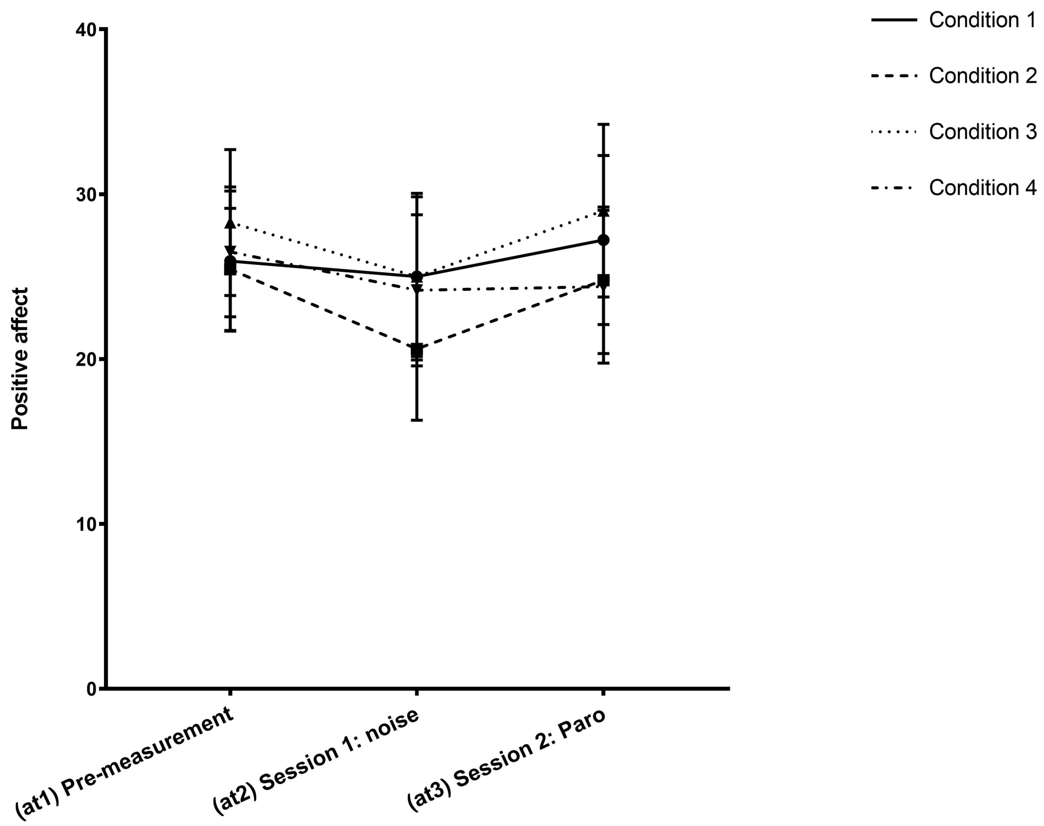


Figure 3.6: The mean with 95% confidence interval for positive affect in PANAS for each condition.

New variables were created in order to assess differences between conditions in terms of changes in positive affect scores. The new variables are $PA\_at2mat1$ = session 1: noise minus pre-measurement, $PA\_at3mat2$ = session 2: Paro minus session 1: noise and $PA\_at3mat1$ = session 2: Paro minus pre-measurement.

There was no significant difference between conditions for $PA\_at2mat1$, $F(3,68) = 1.487, p = 0.226, r = 0.062$. As can be seen in Figure 3.7, it was found that the positive affect scores decreased from the pre-measurement (at1) to the session 1: noise (at2) for all conditions.

There was also no significant difference between conditions for $PA\_at3mat2$, $F(3,68) = 1.675, p = 0.181, r = 0.069$. As shown in Figure 3.7, it was found that the positive affect scores increase from session 1: noise (at2) to session 2: Paro (at3) in all conditions.

There was also no significant difference between conditions for $PA\_at3mat1$, $F(3,68) = 1.264, p = 0.294, r = 0.053$.

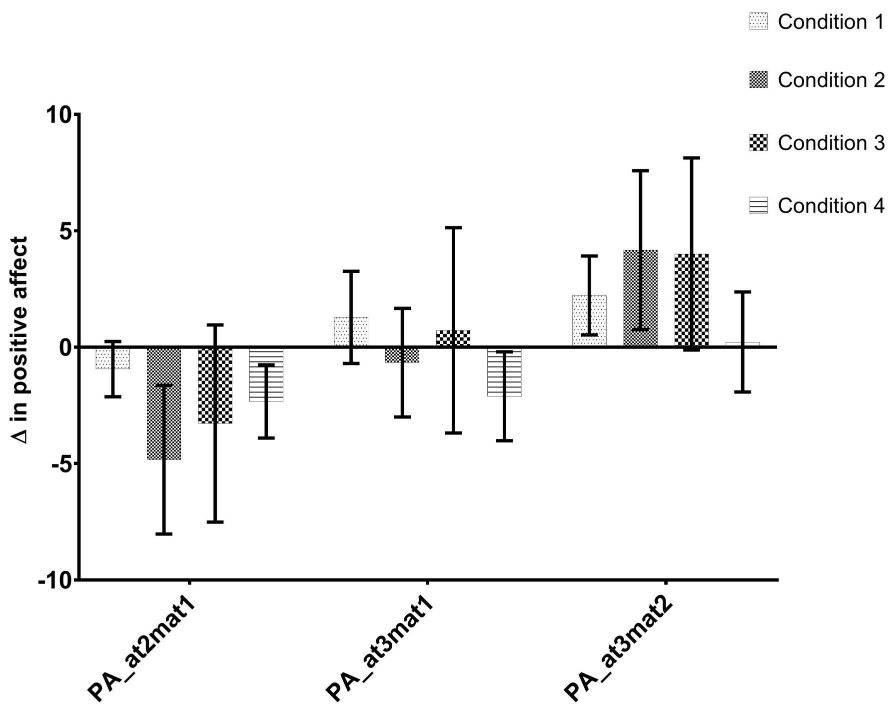


Figure 3.7: The mean with 95% confidence interval for changes in positive affect scores in PANAS for each condition.

Effects of conditions during session 2: Paro

A one-way between-subjects ANOVA showed there was no significant difference between conditions at session 2: Paro on self-reported positive affect, $F(3,68) = 0.881, p = 0.455, r = 0.037$ (Refer to Figure 3.14, page 95).

3.11.4 Negative affect in PANAS

A mixed between-within subjects ANOVA showed there was a significant effect of time on negative affect, $F(2,136) = 34.80, p < 0.0001, r = 0.339$, and no significant effect of conditions,

$F(3,68) = 1.538, p = 0.213, r = 0.064$. There was a significant interaction between condition and time on negative affect, $F(6,136) = 2.361, p = 0.034, r = 0.094$.

Pairwise comparisons between time points were done using Bonferroni correction test. As is apparent in Figure 3.8, the negative affect scores increase significantly after listening to the noise (at2) from pre-measurement (at1), $p < 0.0001$. The negative affect scores also showed significant decrease after the session 2: Paro (at3) from the session 1: noise (at2), $p = 0.002$.

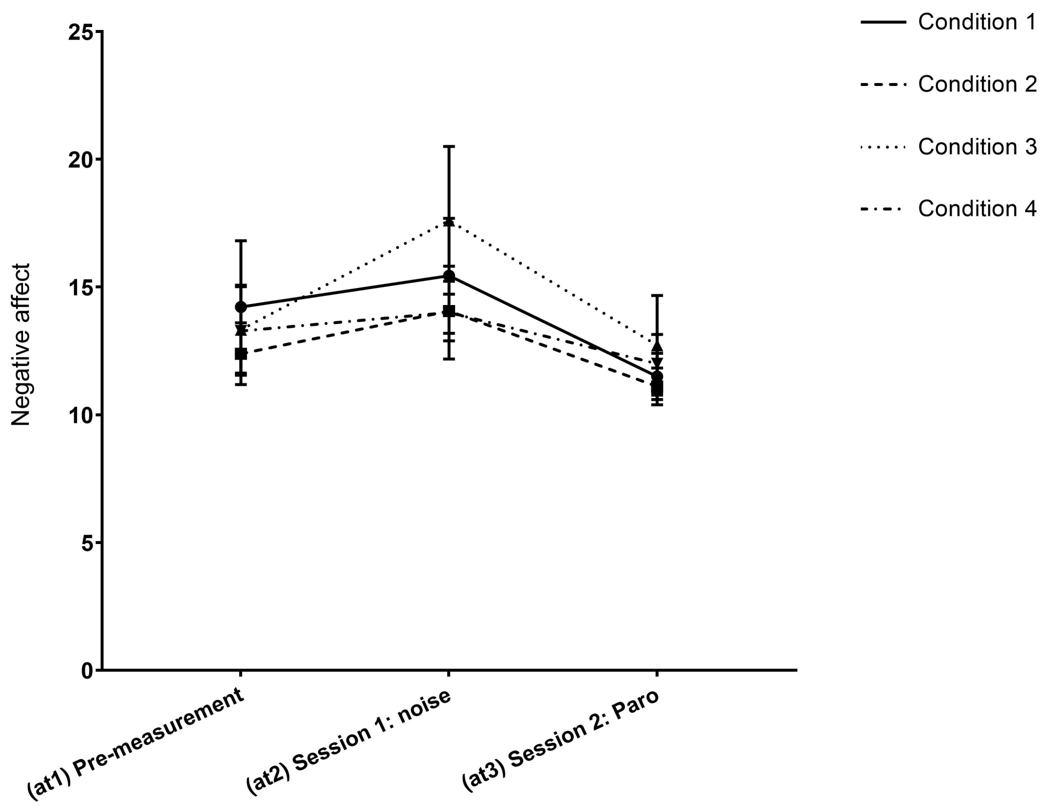


Figure 3.8: The mean with 95% confidence interval for negative affect in PANAS measures to each condition.

New variables were created in order to assess differences between conditions in terms of changes in negative affect scores. The new variables are $NA\_at2mat1$ = session 1: noise minus pre-measurement, $NA\_at3mat1$ = session 2: Paro minus pre-measurement and $NA\_at3mat2$ = session 2: Paro minus session 1: noise.

There was a significant difference between conditions for $NA\_at2mat1$, $F(3,68) = 2.846, p = 0.044, r = 0.112$. However, a post hoc test using Bonferroni correction found no significant difference between conditions. As shown in Figure 3.9, the negative affect scores increase from pre-measurement (at1) to the listening to the noise (at2) for all conditions.

There was a significant difference between conditions for $NA\_at3mat2$, $F(3,68) = 2.864, p = 0.043, r = 0.112$. As evident in Figure 3.9, the negative affect scores decrease from session 1: noise (at2) to session 2: Paro (at3) for all conditions. The decrease in negative affect scores were larger in C3:inactive_stroke than in C4:inactive_rest, $p = 0.043$. There were no significant differences were found between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, C1:active_stroke_talk and C3:inactive_stroke, $p = 1.000$, C1:active_stroke_talk and C4:inactive_rest, $p = 0.399$, C2:active_talk and C3:inactive_stroke, $p = 0.399$, and C2:active_talk and C4:inactive_rest, $p = 1.000$.

There was no significant difference between conditions for $NA\_at3mat1$, $F(3,68) = 1.185, p = 0.322, r = 0.050$. As shown in Figure 3.9, the negative affect scores decrease from pre-measurement (at1) to session 2: Paro (at3) for all conditions.

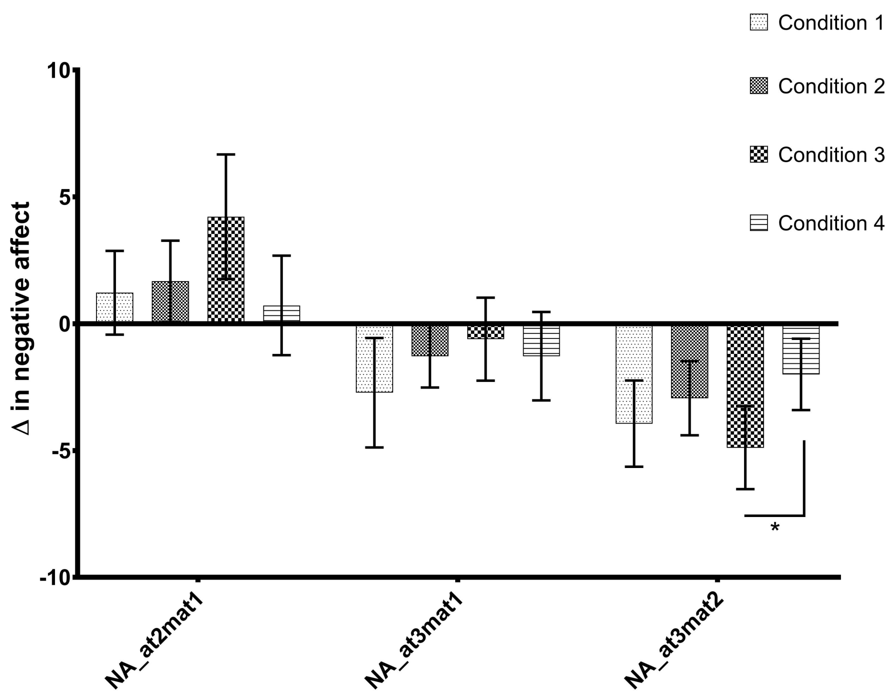


Figure 3.9: The mean with 95% confidence interval for negative affect in PANAS measures to each condition. $*p < 0.05$.

Effects of conditions during session 2: Paro

A one-way between-subjects ANOVA showed there was no significant difference between conditions at session 2: Paro on self-reported negative affect, $F(3,68) = 1.333, p = 0.271, r = 0.056$ (See Figure 3.14, page 95).

3.11.5 Observed facial emotional expressions

Inter-rater reliability A high degree of reliability was found between two evaluators for each emotion across the four conditions, α between 0.980 to 0.988.

Effects of conditions on smile, laugh, disgust and boredom

A one-way between-subjects ANOVA test was used to test for differences between conditions for each coded facial expression: smile, laugh, disgust and boredom. There was a significant difference between conditions in the number of times each of the following expressions occurred: smile $F(3,68) = 12.23, p < 0.0001, r = 0.351$, laugh $F(3,68) = 7.547, p < 0.0001, r = 0.250$, disgust $F(3,68) = 2.914, p = 0.040, r = 0.114$ and boredom $F(3,68) = 14.14, p < 0.0001, r = 0.384$.

For the coded facial expression, smile, a post-hoc comparison found a significant difference between C1:active_stroke_talk and C3:inactive_stroke, $p = 0.009$, between C1:active_stroke_talk and C4:inactive_rest, $p < 0.0001$ and between C2:active_talk and C4:inactive_rest, $p = 0.0001$. The smiling scores were higher in C1:active_stroke_talk compared to C3:inactive_stroke and C4:inactive_rest. The smiling scores also higher in C2:active_talk compared to C4:inactive_rest. (Refer to Figure 3.10). There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 0.726$, C2:active_talk and C3:inactive_stroke, $p = 0.532$ and C3:inactive_stroke and C4:inactive_rest, $p = 0.095$.

For the coded facial expression, laugh, post-hoc comparison found significant difference between C1:active_stroke_talk and C4:inactive_rest, $p = 0.012$, between C2:active_talk and C3:inactive_stroke, $p = 0.012$ and between C2:active_talk and C4:inactive_rest, $p < 0.0001$. The laughing scores were higher in C1:active_stroke_talk compared to C4:inactive_rest. The smiling scores were also higher in C2:active_talk compared to C3:inactive_stroke and C4:inactive_rest. (Refer to Figure 3.10). There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, C1:active_stroke_talk and C3:inactive_stroke, $p = 0.184$ and C3:inactive_stroke and C4:inactive_rest, $p = 1.000$.

For the coded facial expression, disgust, post-hoc comparison found significant difference between C3:inactive_stroke and C4:inactive_rest, $p = 0.039$. The disgust scores were lower in C3:inactive_stroke compare to C4:inactive_rest. (Refer to Figure 3.10). There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, C1:active_stroke_talk and C3:inactive_stroke, $p = 0.196$, C1:active_stroke_talk and C4:inactive_rest, $p = 1.000$, C2:active_talk and C3:inactive_stroke, $p = 0.395$ and

C2:active_talk and C4:inactive_rest, $p = 1.000$.

For the coded facial expression, boredom, post-hoc comparison found significant difference between C1:active_stroke_talk and C3:inactive_stroke, $p < 0.0001$, between C2:active_talk and C3:inactive_stroke, $p < 0.0001$ and between C3:inactive_stroke and C4:inactive_rest, $p < 0.0001$. The boredom scores were lower in C1:active_stroke_talk compared to C3:inactive_stroke. The disgust scores were also lower in C2:active_talk compared to C3:inactive_stroke. However, the boredom scores were higher in C3:inactive_stroke compared to C4:inactive_rest. (Refer to Figure 3.10). There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, C1:active_stroke_talk and C4:inactive_rest, $p = 1.000$ and C2:active_talk and C4:inactive_rest, $p = 1.000$.

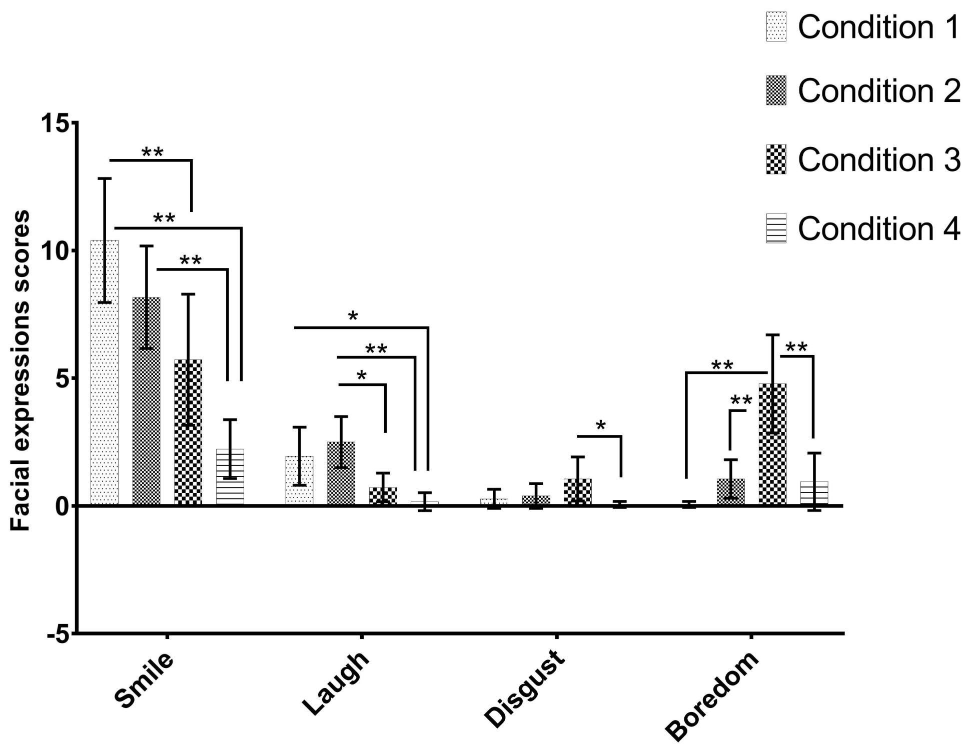


Figure 3.10: The mean with 95% confidence interval for facial emotional expressions to each condition. $*p < 0.05, **p < 0.001$.

Effects of conditions on positive and negative facial expressions:

A further analysis was done to compare scores for positive and negative facial expressions for each condition.

A one-way between-subjects ANOVA test found a significant difference between conditions for positive facial expressions, $F(3,68) = 15.33, p < 0.0001, r = 0.404$. A post hoc test using Bonferroni correction was conducted. There was a significant difference between C1:active_stroke_talk and C3:inactive_stroke, $p < 0.003$, between C1:active_stroke_talk and C4:inactive_rest, $p < 0.0001$ and, between C2:active_talk and C4:inactive_rest, $p < 0.0001$. As shown in Figure 3.11, the positive expressions scores were higher in C1:active_stroke_talk compared to C3:inactive_stroke and C4:inactive_rest. It was found that the positive expressions scores also were higher in C2:active_talk compared to C4:inactive_rest. There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, between C2:active_talk and C3:inactive_stroke, $p = 0.065$ and between C3:inactive_stroke and C4:inactive_rest, $p = 0.086$.

A one-way between-subjects ANOVA test found a significant difference between conditions for negative facial expressions, $F(3,68) = 13.54, p < 0.0001, r = 0.374$. There was a significant difference between C1:active_stroke_talk and C3:inactive_stroke, $p < 0.0001$, between C2:active_talk and C3:inactive_stroke, $p < 0.0001$ and, between C3:inactive_stroke and C4:inactive_rest, $p < 0.0001$. As shown in Figure 3.11, the negative expressions scores were lower in C1:active_stroke_talk compared to C3:inactive_stroke. It was found that the negative expressions scores also were lower in C2:active_talk compared to C3:inactive_stroke. The negative expressions scores were higher in C3:inactive_stroke compared to C4:inactive_rest. There were no significant differences between C1:active_stroke_talk and C2:active_talk, $p = 1.000$, between C1:active_stroke_talk and C4:inactive_rest, $p = 1.000$ and between C2:active_talk and C4:inactive_rest, $p = 1.000$.

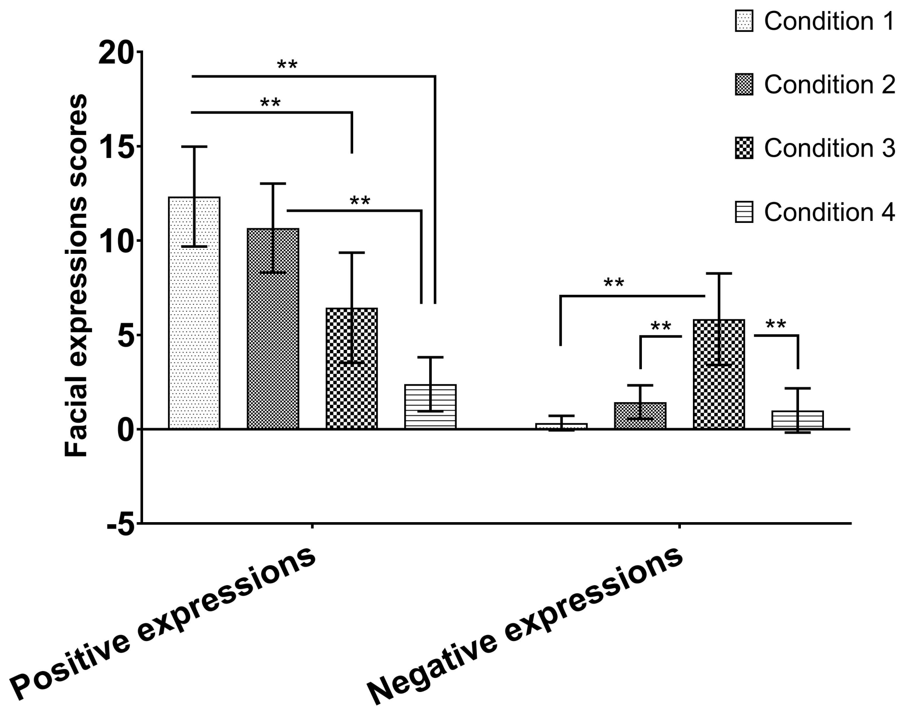


Figure 3.11: The mean with 95% confidence interval for negative and positive expressions to each condition. $**p < 0.001$.

3.11.6 Skin conductance responses

A mixed between-within subjects analysis of variance showed there was a significant effect of time on skin conductance response (SCR), $F(2.508, 170.5) = 83.22, p < 0.0001, r = 0.550$, and no significant effect of conditions, $F(3, 68) = 0.901, p = 0.445, r = 0.038$. There was significant interaction between condition and time on SCR, $F(7.523, 170.5) = 11.64, p < 0.0001, r = 0.339$.

Pairwise comparisons between time points were done using Bonferroni correction test. The SCR increased after listening to the noise (t2) from pre-measurement (t1), $p < 0.0001$, from the session 1: noise (t2) to session 2: Paro (t3), $p < 0.0001$. (Refer to Figure 3.12). However, no significant between time-points baseline (t1) and post-measurement (t4), $p = 0.028$. Then the SCR showed a significant decrease from session 2: Paro (t3) to post-measurement (t4), $p < 0.0001$. (Refer to Figure 3.12).

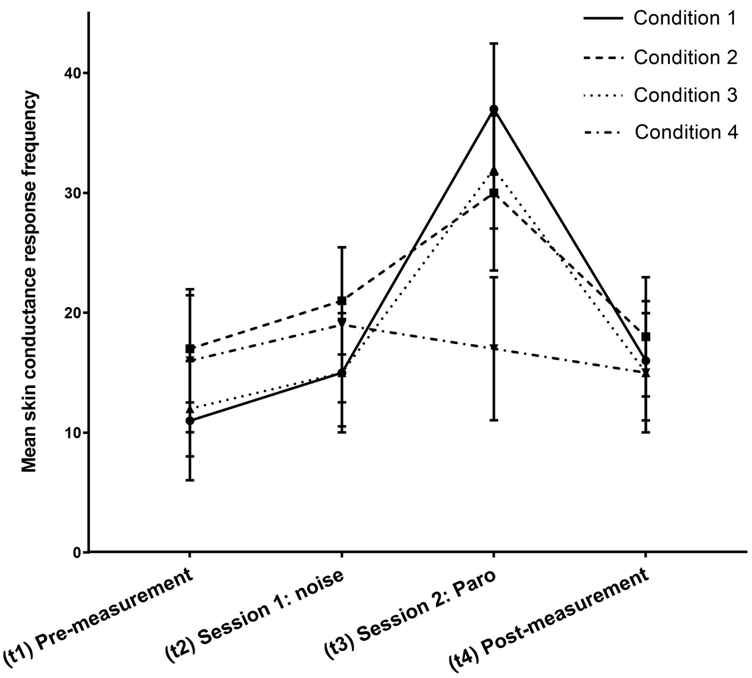


Figure 3.12: The mean with 95% confidence interval for skin conductance responses to each condition.

New variables were created in order to check if there were differences between conditions in terms of changes in SCR. The new variables are SCR_t2mt1 = session 1: noise minus pre-measurement, SCR_t3mt2 = session 2: Paro minus session 1: noise, SCR_t4mt3 = post-measurement minus session 2: Paro, and SCR_t4mt1 = post-measurement minus pre-measurement.

There was no significant difference between conditions for SCR_t2mt1, $F(3,68) = 0.031, p = 0.993, r = 0.001$. As is apparent in Figure 3.13, the SCR increased from pre-measurement (t1) to session 1: noise (t2) for all conditions.

There was a significant difference between conditions for SCR_t3mt2, $F(3,68) = 23.62, p < 0.0001, r = 0.510$. As is shown in Figure 3.13, the increase in SCR were significantly higher in C1:active_stroke_talk compared to C2:active_talk, $p < 0.0001$, in C1:active_stroke_talk compared to C4:inactive_rest, $p < 0.0001$, in C2:active_talk compared to C4:inactive_rest, $p = 0.006$ and in C3:inactive_stroke compared to C4:inactive_rest,

$p < 0.0001$. No significant difference was found between C1:active_stroke_talk and C3:inactive_stroke, $p = 0.3920$, and C2:active_talk and C3:inactive_stroke, $p = 0.066$

There was a significant difference between conditions for SCR_t4mt3, $F(3,68) = 12.81, p < 0.0001, r = 0.361$. As is evident in Figure 3.13, the decrease in SCR were significantly greater in C1:active_stroke_talk compared to C2:active_talk, $p < 0.021$, in C1:active_stroke_talk compared to C4:inactive_rest, $p < 0.0001$, in C2:active_talk compared to C4:inactive_rest, $p = 0.043$ and in C3:inactive_stroke compared to C4:inactive_rest, $p < 0.0001$. No significant difference was found between C1:active_stroke_talk and C3:inactive_stroke, $p = 1.000$, and C2:active_talk and C3:inactive_stroke, $p = 0.382$.

There was no significant difference between conditions for SCR_t4mt1, $F(3,68) = 1.681, p = 0.179, r = 0.069$.

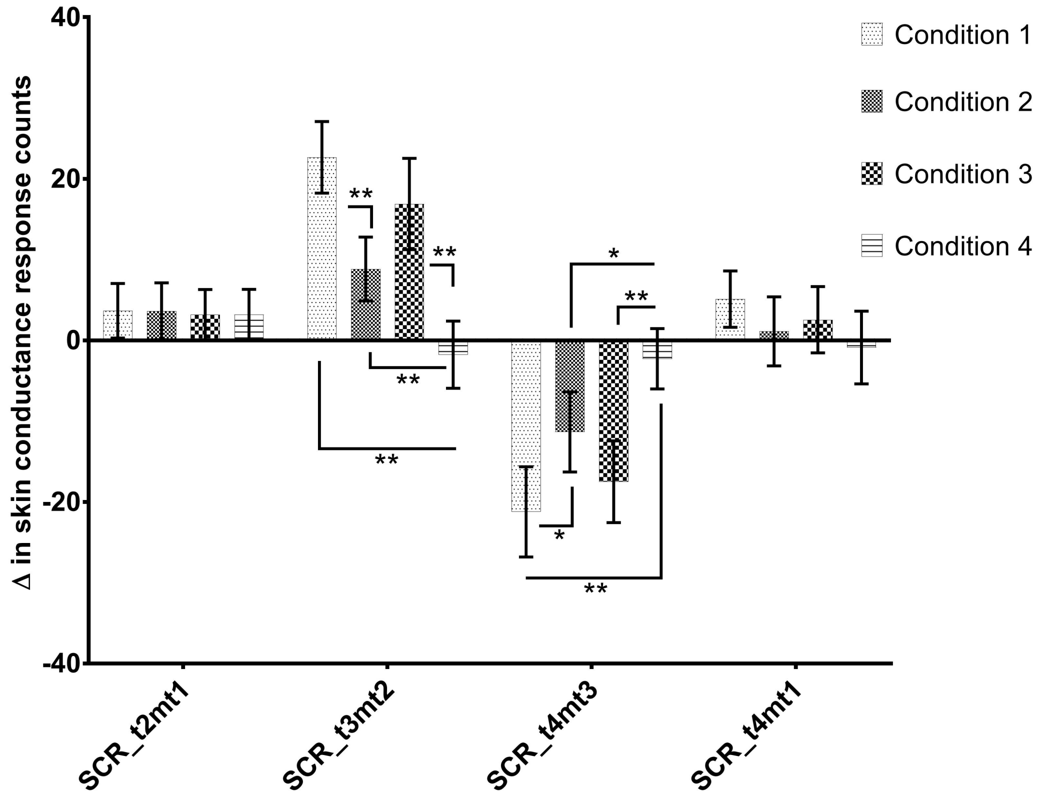


Figure 3.13: The mean with 95% confidence interval for changes scores in skin conductance responses to each condition. $*p < 0.05, **p < 0.001$.

Effects of conditions during session 2: Paro

A one-way between-subjects ANOVA showed there was a significant difference between conditions at session 2: Paro on skin conductance responses, $F(3,68) = 10.03, p < 0.0001, r = 0.307$. SCR were significantly higher in C1:active_stroke_talk, $p < 0.0001$, C2:active_talk, $p = 0.010$, C3:inactive_stroke, $p = 0.001$ compared to C4:inactive_rest, as can be seen in Figure 3.14. No significant difference was found between C1:active_stroke_talk and C2:active_talk, $p = 0.303$, C1:active_stroke_talk and C3:inactive_stroke, $p = 1.000$, and C2:active_talk and C3:inactive_stroke, $p = 1.000$.

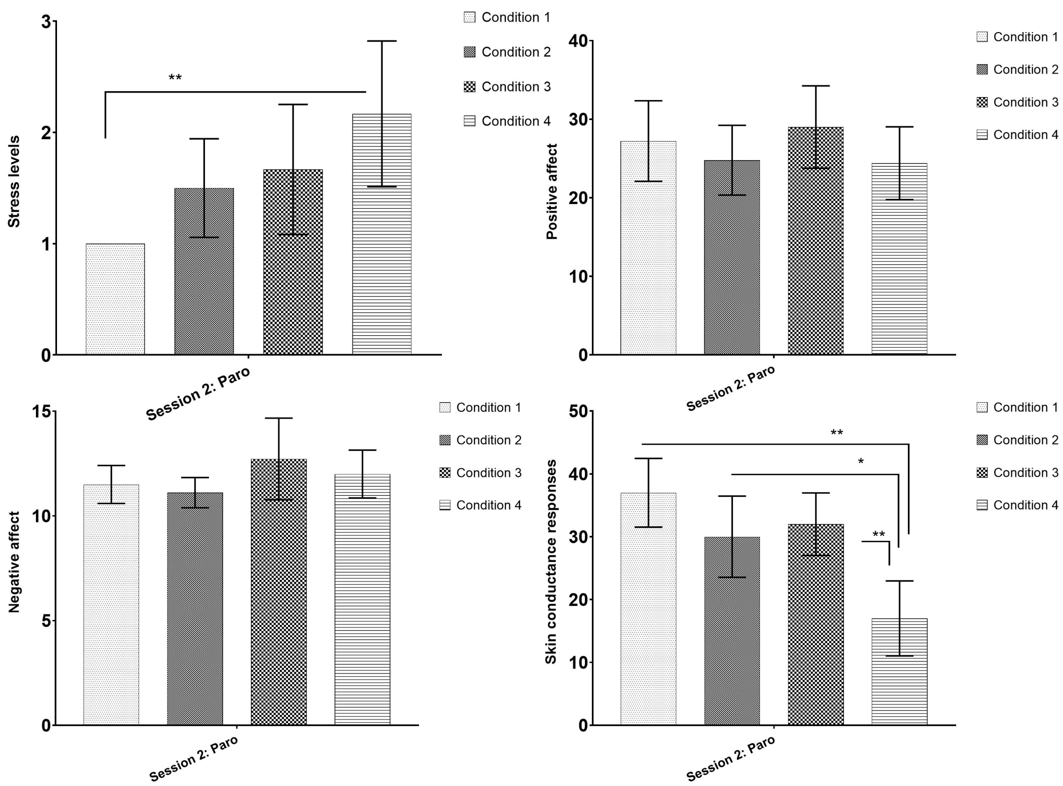


Figure 3.14: The mean with 95% confidence interval for self-reported stress, positive, negative and skin conductance responses to each condition at session 2: Paro. $*p < 0.05, **p < 0.001$.

3.12 Discussion

Aims of Study 1: The primary aims in Study 1 were to explore the effects of short-term interactions with Paro, and the effects of the short-term interaction on stress responses. Additionally, in order to understand the effects in depth, Study 1 examined the effects of different human and robot characteristics that could benefit psychophysiology. Study 1 compared the effects of characteristics such as talking and stroking with the robot, and the embodiment of the robot (i.e. soft, fur, cute face, sound, body movement). The effects of Paro on humans in this study were explored using observed facial expression, self-reported stress, positive and negative affect and physiological responses (i.e. skin conductance responses).

Findings: The discussion of the results were divided into two parts: in Part 1, the effects of conditions during intervention session, and in Part 2, the effects of interaction with the robot on stress responses.

3.12.1 Part 1: Effects of conditions during intervention session

In this Part 1, there will be discussion of the findings on self-reported and skin conductance responses, and observed facial expression during the intervention session.

Self-reported responses:

In Study 1, it was found that participants who had full interaction with an active Paro (C1:active_stroke_talk) had significantly lower self-reported stress compared to no interaction with Paro (C4:inactive_rest), *partially supporting H1(a)*. From this finding, it was suggested that interactive intervention with Paro (C1:active_stroke_talk) was more effective at lowering stress than resting in the presence of an inactive Paro (C4:inactive_rest). The results are consistent with the study by Saito et al. (2003)[4]. They found that the levels of stress were higher

[4] Study on the effects of interacting with Paro on stress hormones, page 26

in a condition where no Paro was introduced compared to both interactive Paro conditions. However, in terms of reducing stress, Saito et al. (2003) found that active Paro had a greater impact on stress compared to inactive Paro. These findings differ from those of results of Study 1 found at session 2: Paro, where no difference was found between interacting with an active Paro and stroking an inactive Paro, as measured by self-reported stress at session 2: Paro.

No differences were also found in self-reported positive and negative affect across conditions during session 2: Paro. However, the results of self-reported stress, positive and negative affect in Study 1 are in line with the work by (Saito et al. 2003, Wada et al. 2004*a*), who found also no difference between conditions that involved interacting with active or less active Paro in mood improvement (Saito et al. 2003), and in stress reduction (Wada et al. 2004*a*).

In the study by Wada et al. (2004*a*), the interactions with Paro happened in groups, so the results may have been influenced by interactions with other people in the groups. In the study by Saito et al. (2003) suggested that the presence of Paro in the ward may have influenced the improvement of moods, regardless of Paro behaviours (i.e. active or less active). *Therefore, the recent findings from Study 1 showed that the effects of Paro are also beneficial to people that interact with the pet robot individually, and even when the Paro has been turned off.*

Skin conductance response:
In the case of skin conductance response (SCR), *H1(b) is partially supported*, Study 1 found interaction sessions with Paro in three conditions (C1:active_stroke_talk, C2:active_talk and C3:inactive_stroke) resulted in significantly higher SCR than C4:inactive_rest. The findings in Study 1 are consistent with past research of Paro, where it was suggested that interacting with active Paro increased the skin conductance responses (Mitsui et al. 2001*a*, Petersen et al. 2017). Despite higher levels of SCR in both of these studies, interestingly, participants showed an improvement in moods (Mitsui et al. 2001*a*), and an improvement in observed scales of stress and anxiety (Petersen et al. 2017). *This suggests that the higher SCR during interaction with the pet robot is related to positive arousal.*

In addition, the study by Mitsui et al. (2001*a*) found an increased SCR from baseline to experimental conditions session across conditions. But, the authors also found that the SCR were significantly lower in the inactive Paro condition than in the active Paro. These findings are different to the results of Study 1, where *the increased SCR were higher in all interactive conditions (C1:active_stroke_talk, C2:active_talk and C3:inactive_stroke) compared to C4:inactive_rest, where no interaction with inactive Paro was allowed.*

Study 1 suggests that the higher SCR is due to stroking a soft object (C1:active_stroke_talk and C3:inactive_stroke), and talking with Paro (C1:active_stroke_talk and C2:active_talk). However, when comparing between stroking or/and talking, C1:active_stroke_talk and C3:inactive_stroke results in a higher SCR compared to C2:active_talk. Although the difference between C2:active_talk and C3:inactive_stroke is not significant, this indicates that stroking Paro showed a higher SCR arousal compared to talking to Paro (C3:inactive_stroke). As mentioned in Chatel-Goldman et al. (2014), hand stroking in tactile interaction increases SCR, but also alleviates stress responses (Remington 2002a, Cramer et al. 2009, Joranson et al. 2015). *The stress-reducing effects in Study 1 could be seen in the low self-reported stress and negative responses, despite the high SCR during session 2: Paro in C1:active_stroke_talk, C2:active_talk and C3:inactive_stroke.*

Observed facial expressions:

Other notable findings in Study 1 are the significantly higher frequency of expressions of disgust and boredom observed in C3:inactive_stroke (stroking inactive Paro) than in C4:inactive_rest (no interaction with Paro). Additionally, lower frequency of negative facial expressions was observed in conditions with an active Paro (C1:active_stroke_talk and C2:active_talk). These results *partially supported H1(c)*. These effects are similar to a previous study by Takayanagi et al. (2014), where they found that participants showed more positive emotional expressions when interacting with an active Paro compared to inactive object. In addition, Robinson et al. (2013a) and Moyle et al. (2017a) found that patients tended to smile more when interacting with Paro compared to a non-animal robotic companion, and an inactive toy. In term of the perceptions of family members, they mentioned that playing with a responsive Paro benefited the patients in terms of increased social interaction and mood improvements (Moyle et al. 2017b). Further findings of increased in positive facial expressions after interacting with Paro were also found in studies by Wada et al. (2004*a*) and Moyle et al. (2017b).

3.12.2 Part 2: Effects of interaction with the robot on stress responses

Part 2 discussed the findings on the stress-reducing effects of Paro on induced stress.

It was found that when comparing changes in scores from baseline to final measurements, no further reduction in the self-reported stress and skin conductance response (SCR) was found in all conditions. Study 1 did not find a greater improvement in positive affect at the end of experiment, but the PA returned to baseline at the end of experiment in all conditions. This suggests that even the PA decrease from baseline after the stress session ended, but the Paro conditions help to elicit the PA again. Interestingly, only self-reported negative affect showed a significant decrease below the baseline at the end of experiment in any of the conditions. This suggests that a session with Paro, whether the robot is active or inactive, and whether there was or was no any interaction with the robot, results in a decrease in negative moods. Therefore, it is acknowledged that most of the findings in Study 1 do not support the idea that different characteristics of Paro and human behaviour in short-term session have different effects on the stress responses. Therefore the final *hypothesis (H2) is not supported*.

Skin conductance responses:
It is worth noting that, the skin conductance response (SCR) returned to baseline for all conditions at the end of experiment. This may suggests an increase of parasympathetic activity and decrease in sympathetic activity (Pour et al. 2010), a detail about SCR in section 2.6.4 on page 43.

However, it is also possible that the absence of any further reduction in SCR at the end of experiment compared to baseline may be due to the restrictions imposed about how to interact with Paro. In Study 1, for example in C1:active_stroke_talk, participants were asked to only stroke and talk with the Paro, but they were not allowed to pick up and hug the Paro. Hugging the Paro while petting it might have had a greater effect in reducing the stress. Furthermore, a study by Etzi et al. (2018) has previously shown that lower arousal of SCR is associated with slow and gentle hand movements (i.e. stroking and touch). However, it was beyond the scope of Study 1 to look at the effects of different levels of stroking and hugging the Paro during the robot therapy intervention.

The findings of Study 1 are in contrast with the study by Mitsui et al. (2001*b*). In Study 1, the SCR was increased significantly from baseline to session 1: noise for all conditions, and from session 1: noise to session 2: Paro (except C4:inactive_rest). But, the SCR was significantly higher in session 2: Paro (at t3) compared to session 1: noise (at t2). In Mitsui et al. (2001*b*) study, the authors found an increase in SCR from baseline to a mental task session, and from baseline to a session with Paro (either active or inactive). But, there was a greater increase in SCR for mental task session compared to Paro sessions. There was also a trend showing a greater increase in SCR when stroking with an inactive Paro compared to active Paro.

Mitsui et al. (2001*b*) suggested that the increased SCR during the mental task was due to concentration when performing the task. It was also suggested that the increased SCR during inactive Paro session was because the participants did not like the inactive Paro. However, no statistical comparison was reported between the two sessions (Paro versus mental stress). As the study tended to investigate changes in physiological responses from baseline to a different session (i.e. movement of hand and mental task), and the number of participants was small, so the study could be underpowered.

Another possible reason for the contradictory results of SCR from past study by Mitsui et al. (2001*b*) could be because a different stressor was used to induce the stress in Study 1.

However, past research showed noise exposure increased negative affect and skin conductance responses (Yang et al. 2014, Cvijanović et al. 2017). These results were supported by Study 1 with the increases in SCR during the noise exposure, and increases in self-reported stress and negative affect after listening to noise. Nonetheless, despite the reliability of the noise to induce mild stress, *the noise might be not as too stressful when as to stress induced by the mental task* used in Mitsui et al. (2001*b*) study. This could explain why in Study 1 the SCR arousal in session 1: noise was lower than in session 2: Paro.

Another suggestion, as this is a first study of Paro that used induced noise as a stressor, *it may be that a 3 minutes session with Paro is too short, and not enough to reduce the effects of previous stress, hence the increased SCR*.

Even though the findings in Study 1 showed a contrasting result from Mitsui et al. (2001*b*) in terms of SCR arousal between interaction in active and inactive Paro, it should be noted that the study by Mitsui et al. (2001*b*) used a within-subject design and a small number of participants. One of the disadvantages of the within-subject design, the participants could be influenced by personal bias effect.

But, there are some findings in Study 1 that are consistent with Mitsui et al. (2001*b*), since results of Study 1 suggested that physical interaction with Paro was associated with high arousal of SCR.

Talking versus stroking:

Study 1 also includes a new finding based on an interaction with Paro without touching it, but only talking to it (C2:active_talk). Many past studies of Paro found an improvement in moods and physiological responses after a session with Paro, where participant got to talk to and stroke it (Robinson et al. 2013c, 2015). However, less is known about whether only talking to Paro could help in reducing stress and improving mood. The results of Study 1 suggest that talking to Paro also helps increase positive moods and decrease stress and negative moods, based on changes in scores between session 1: noise and session 2: Paro (at3mt2).

However, there was a slight decrease trend in PA in C2:active_talk, and an increased trend in PA in C1:active_stroke_talk and C3:inactive_stroke, at the end of experiment compared to baseline (at3mat1). *This may suggests that stroking is one of the factors that have a strong influence on positive moods compared to talking to Paro.* This result is consistent with a study by (Vormbrock & Grossberg 1988), where the authors found that *there was an improvement in physiological responses when stroking a dog compared to just talking to a dog.* The results of Study 1 are compared here with research in animal therapy, as there has been no previous studies of Paro that investigating the effect of stroking compared to talking on stress responses.

Table 3.2 showed a summary of findings of Study 1 based on hypothesis testing.

Table 3.2: Findings of Study 1 based on hypothesis tested.

Hypothesis	Supported?	Responses	Finding	Refer to
H1(a)	Partially	Self-reports	The self-reported stress was significantly lower when interacting with Paro (C1:active_stroke_talk) than in the no interaction condition (C4:inactive_rest). High self-reported positive affect and low negative affect were found across all conditions.	Figure 3.14, page 95
H1(b)	Partially	Physiology	SCR in full interaction with Paro (C1:active_stroke_talk) was higher when compared to rest condition (C4:inactive_rest), but the SCR was also high in two other conditions (C2:active_talk and C3:inactive_stroke).	Figure 3.14, page 95
H1(c)	Partially	Observed facial expression	Interacting with both active Paro conditions showed more positive facial expressions than in C3:inactive_stroke and C4:inactive_rest. Negative expressions were lower in C1:active_stroke_talk and C2:active_talk than in the other two conditions.	Figure 3.11, page 92
H2(a)	No	Self-reports	No significant effect of condition on self-reported stress, positive and negative affect.	Figure 3.5, page 82, 3.7, page 85 and 3.9, page 88

Table 3.2 continued from previous page

Hypothesis	Supported?	Responses	Finding	Refer to
H2(b)	No	Physiology	No significant effect of condition on skin conductance responses. All conditions returned to baseline at the end of experiment.	Figure 3.13, page 94

3.13 Strengths of Study 1

Study 1 removed some of the limitations of previous studies by using a between-subject and randomized study design, in order to reduce bias. This study was also done in controlled settings, in order to eliminate the effect of the presence of a researcher during intervention session with a robot, in order to reduce the mediator effect (Robinson et al. 2013c). This helps in understanding the sole effects of Paro on reducing stress without external variables that could confound the results. Furthermore, the number of participants were higher (n = 76) compared to previous studies of Paro (n less than 10) that looked at physiological improvement in healthy young adults (Mitsui et al. 2001*b*, Saito et al. 2003).

This is also the first study of Paro that looks at the effects of Paro on induced environmental stress. The reduction in self-reported negative affect to baseline at the end of experiment suggested that Paro may be effective at reducing the stress to this type of stressor. The positive interactions during session 2: Paro also could provide the benefits of reducing the risk of cardiovascular diseases that can occur when people are exposed to noise.

3.14 Important considerations for the next study

3.14.1 Influence of stroking pet robot on self-reported responses

As no significant differences were found in Study 1 between C1:active_stroke_talk and C3:inactive_stroke, it is suggested that stroking the soft fur of Paro is one of the factors that could contribute to stress reduction, as reported in self-reported stress. This leads to questions about whether stroking soft fur alone could help in reducing stress.

It was further suggested that the appearance of an inactive Paro in condition 3 and 4 also may have influenced the reduction in self-reported negative affect at the end of experiment compared to baseline. An increased trend in positive affect was also found at the end of experiment in C3:inactive_stroke. Therefore, in order to make a clear conclusion about the effects of stroking during interaction and on stress induction (i.e. noise), the next study will eliminate the presence of an inactive Paro. The effects of stroking on stress reduction will be further investigated with a comparison between an active Paro and an inactive object -a bolster covered with soft fur.

3.14.2 Interpreting the effects of pet robot therapy on physiological responses

Regarding the skin conductance response (SCR), there are a number of possible reasons for the increased SCR during session 2: Paro in Study 1:

Novelty effect: The increase in skin conductance responses could be related to the excitement and intensity of participants during interaction with Paro. As noted earlier, this reason was supported by the results of Study 1, and a previous study which found that interaction with Paro made people overstimulated (Jung et al. 2017). The overstimulation could be due to the novelty effect of interacting with a new thing. Adequate time to become familiar with the robot before the intervention could reduce the novelty effect of Paro.

Familiarisation: Apart from being over stimulated by interacting with Paro, participants may have found interacting with the Paro stressful because it was new to them. In Study 1, participants were only given a brief instruction about how to interact with the Paro (e.g. stroking and talking to Paro) and how the Paro works. They were not given any time to become familiar with the Paro before the session with it. Previous studies have mentioned how familiarisation with a pet will affect the stress reduction (Herzog 2011, Saunders et al. 2017). For instance, a comparison study between pet owners and non-pet owners showed that participants who were accompanied by their own pet during the stressful event showed a greater stress reduction (Barker et al. 2010). Therefore, the next study will look at the effects of some familiarisation time before the intervention with the Paro.

Additional noises: Furthermore, it was found that the skin conductance response (SCR) arousal during session 1: noise was less than SCR found in with session 2: Paro. The changes in arousal to the stressor will be further explored by adding a new noise (babies crying) to the current noise. As the participants recruited were young adults in a student population, the new noise might be unfamiliar to young adults, and should result in increased stress by inducing high arousal SCR. Past research also found that listening to a baby crying affected the sympathetic nervous system activity by increasing the SCR and heart rate (HR) (Frodi et al. 1978, Casanova et al. 1992, McCanne & Hagstro m 1996, Groh & Roisman 2009, Out et al. 2010). The physiological changes were also found to be higher in an unfamiliar noise condition compared to a familiar noise condition (Kjellberg et al. 1996). Therefore, in next study, a new noise will be used, that is a combination of sound of heavy traffic, crowd conversation and babies crying.

Length of session with Paro: Additionally, a longer session with Paro may also have influenced the amount of stress reduction to the induced stress. It has been shown that even durations as short as 2 to 5 minutes is enough to stimulate positive interaction with pet robot (Kramer et al. 2009, Kawaguchi et al. 2012), and this is also supported from the results of Study 1, where there was an increase in positive moods. Therefore, in the next study, participants will have a 5 minutes session of intervention session.

Preventive stress intervention: Additionally, the effects of stress-reducing Paro will be further investigated by changing the order of stressor and intervention session with a robot. In the next study, participants will interact with Paro before being exposed to induced stress. In past research, Paro has been used as a preventive of long-term stressors such as reducing agitation in people with Dementia (Pedersen 2011, Takayanagi et al. 2014, Joranson et al. 2015, Thodberg et al. 2016, Moyle et al. 2017a). However, identifying those effects in short-term interaction could provide an immediate answer about which interventions may have a greater improvement in moods and reduce acute stress. Furthermore, this intervention setting should give a clearer picture of the effects of interacting with Paro on skin conductance response (SCR) without the possible contaminating effect of prior induced stress i.e. noise.

Physiological measurements: In the next study, different physiological measures will also be used to measure the effects of Paro: heart rate and heart rate variability. Past research of the effects of Paro on physiological measures suggested that different physiological measures appear to show different responses to interaction with Paro (Mitsui et al. 2001*b*, Wada et al. 2005b, Carretero & Sáez 2017, Petersen et al. 2017). This should help to provide a clearer evidence about the stress-reducing effects of Paro in terms of physiological effects.

3.15 Conclusion

Overall, this study found that there was a reduction in negative moods at the end of the experiment compared to the baseline. The findings of Study 1 suggest that interactivity, responsiveness, appearance and the soft texture of Paro are all associated with improvements in negative moods. It was also found that interacting with an active Paro can stimulate more positive facial expressions compared to an inactive Paro. Study 1 also pioneered the investigation of the effects of Paro robot intervention on environmental stress. All these findings help to provide an better understanding of the characteristics of the robot and humans that may be able to reduce the effects of induced stress. This helps to highlight a new perspective on the stress-reducing effects of Paro on physical or mental health that affected by noise exposure.

Chapter 4

Study 2: Comparing effects of interacting with an active Paro robot and an inactive bolster as pre-stress intervention

4.1 Summary of Study 1

The main aims of Study 1 were to investigate the effects of Paro during short-term interaction and on noise-induced stress. A further aim was to look at the effects of different characteristics of Paro such as movement, voice and it being presented as an inactive object (e.g. switched off), and human behaviour towards robot such as talking and stroking.

The findings from Study 1 indicate that when participants fully interacted with an active Paro, they had lower self-reported stress compared to a condition in which the participants were accompanied by an inactive Paro (during intervention session: a comparison between Condition 1 (C1:active_stroke_talk) and Condition 4 (C4:inactive_rest)). Furthermore, more positive facial expressions were observed during participants' interaction with an active Paro. Physiological responses recorded during the interaction with Paro in three conditions (except C4:inactive_rest) also showed an increase in skin conductance responses (SCR) from baseline, which is assumed to be related to positive experience, as indicated by self-reported positive affect. Although there was a lack of significant differences between conditions at the end of the experiment, interestingly, the self-reported negative affect was reduced further when compared to baseline. These findings suggested that even the mere presence of Paro following noise exposure may help in alleviating the effects of stress by reducing negative moods. Another suggestion, is that also possible that participants had their ability to cope with the stress responses without any intervention.

It was further suggested in Study 1 that the lack of familiarity, stroking or the novelty effect of the Paro may have resulted in increased skin conductance responses. Based on the reduction in self-reported negative affect after interacting with Paro in all conditions, it was also questioned whether prior interaction with Paro would have similar stress-reducing effects on subsequent stress induction.

4.2 Current study: Study 2

Study 2 aims to better understand the effects of Paro, and its effect on stress responses. Study 2 was undertaken in part to explore the effects of familiarity with an active Paro. Study 2 also compared the effects of interacting with active Paro and stroking a bolster covered with soft fur.

It is also possible that Paro might result in a reduced response to stress inducing noise if the participants spent time interacting with it before they were exposed to the stress inducing noise. So, in Study 2, participants interacted with the Paro before the stress inducing noise. A new sound was added to the previous noise used in Study 1 i.e. babies crying.

The effects were again measured using self-reports and physiological responses, and observed facial expressions. Additionally, in order to understand further the effects of physiological responses to Paro or stress, two physiological measurements were added together with skin conductance responses namely heart rate (HR) and heart rate variability (HRV): standard of all normal RR intervals (SDNN) and high-frequency (HF).

To sum up, the new things explored in Study 2, are as followed:

1. Comparing the effects of interacting with a Paro robot with and without familiarity.
2. Comparing the effect of stroking a bolster and interacting with a Paro robot.
3. Does earlier interaction with a Paro or stroking a bolster reduce the stressful effects of noise?
4. Examining the effects created during experiment on other physiological responses i.e. heart rate and heart rate variability.

Each of these new points is explained further in the next section.

4.3 Terms

The terms listed below are defined for the scope of Study 2:

- *Time-point* refers to the time when the physiological measure was taken. There are four time-points in this study where time-point 1 labelled as t1, time-point 2 labelled as t2, time-point 3 labelled as t3 and time-point 4 labelled as t4.
- *After time-point* refers to the time when the participants need to fill in the questionnaires. Participants fill in the questionnaires after each physiological measures time-point. There are four after time-points in this study where after time-point 1 labelled as at1, after time-point 2 labelled as at2, after time-point 3 labelled as at3 and after time-point 4 labelled as at4.
- *Session 1* refers to intervention session at time-point 2. There are four conditions in this study where Condition 1 labelled as C1, Condition 2 labelled as C2, Condition 3 labelled as C3, and Condition 4 labelled as C4.

- *Session 2* refers to a stress session at time-point 3, where the participants were asked to listen a noise.

- *Changes in scores* refers to new variables created for changes in scores between time-points. The variables for self-reported responses labelled as at2mat1 (changes in scores between session 1 and pre-measurement), at3mat1 (changes in scores between session 2 and pre-measurement), and at3mat2 (changes in scores between session 2 and session 1). The variables for physiological responses labelled as t2mt1 (changes in scores between session 1 and pre-measurement), t3mt2 (changes in scores between session 2 and session 1), 4mt3 (changes in scores between post-measurement and session 2), and t4mt1 (changes in scores between post-measurement and pre-measurement).

4.4 Rationale of Study 2

The following are the rationale of Study 2:

4.4.1 Familiarity:

Study 2 was designed to investigate the effect of an opportunity to get used to the Paro before the experiment begins. The effect of familiarisation was assessed by including a condition in which participants were explicitly taught how to interact with the Paro before the start of the experiment. Previous studies of animal therapy mentioned the stress-reducing effects associated with familiarisation with a pet. For instance, pet owners would experience lower stress compared to participants who do not own a pet or who played with an unfamiliar pet (O'Haire 2010, Beck 2014). In study by Okita (2013), the authors found that no reduction in anxiety level after interact with Paro as the results of not being familiar with Paro, hence increased their stress.

Based on the results from the skin conductance responses in Study 1 on page 105, it was suggested that the SCR levels may have been high because participants were not familiar with the Paro robot. Therefore, Study 2 is designed to further understand the factors that might have influenced the high arousal in SCR, hence the effects of familiarisation with Paro will be explored.

It was anticipated that the familiarity with Paro before the intervention session might reduce the SCR and self-reported stress related responses, and induce positive moods compare to those who meeting the Paro for the first time, or do not interact with Paro.

4.4.2 Effect of stroking:

Study 2 was also designed to determine to what extent stroking an active Paro will affect the self-reports and physiological responses compared to stroking a soft object (i.e. a bolster). This is because in Study 1, Paro was always present in all conditions, therefore the effects of stroking itself cannot be determined because it might have been influenced by the cuteness of Paro.

Past research mentioned how touching and stroking a pet robot resulted in a positive effect on psychological and physiological responses (Mitsui et al. 2001*a*, Bemelmans et al. 2012, Pu et al. 2018). One study of Paro by Mitsui et al. (2001*a*), compared interaction with Paro and an object of a similar size to Paro, and found an increase in heart rate and respiration when the participants interacted with an active Paro, but improved the moods. But, the study used a within-subject design which meant that participants' responses in one condition might have affected other conditions as pointed out by Bordens & Abbott (2002). It is also difficult to make a strong conclusion about how the Paro affected the physiological responses in the study reported by Mitsui et al. (2001*a*), because their sample size was small (n = 6).

Other factors that might contribute to the positive effects is the fur. The softness of fur induces pleasure when interacting with an animal (Kobayashi et al. 2017). These effects also can be related to touch or physical contact with soft fur, as touch is one of the natural ways of interacting or communicating with a pet robot.

Therefore in Study 2, *it was anticipated that interacting with Paro might reduce the stress responses and induce positive responses compare to stroking a bolster.*

4.4.3 Preventive stress intervention:

Additionally, in Study 2, the effect of interacting with Paro or a bolster on a subsequent stress-inducing session was examined. In Study 1, the noise session preceded the Paro session. However, in Study 2, the noise session followed the Paro or bolster session. This enables the assessment of the extent to which a prior session with a Paro could reduce the stressful effects of subsequently encountered noise.

It was anticipated that interacting with Paro might reduce the stress responses and induce positive responses compare to stroking a bolster.

Summary of method of Study 2:

Study 2 had four time points of 5 minutes measurement. A pre-measurement/baseline (t1) was followed by 5 minutes of Paro/bolster/rest session (t2), a 5 minutes noise session (t3) and a 5 minutes post-measurement (t4). Participants were assigned randomly to one of four conditions. In condition 1 (C1), the experimenter gave them 2 minutes to become familiar with and play with a Paro robot before the experimenter started collecting their responses. Then they spent time with the robot again for 5 minutes. In condition 2 - 4, participants did not have a chance to get familiar with the Paro robot but were given a neutral article to read for 2 minutes. In condition 2 (C2), participants were given 5 minutes to play with the robot during the robot session. In condition 3 (C3), participants did not have a session with the robot but were asked to stroke a bolster. The bolster had a similar soft furry covering to that of the Paro robot. In Condition 4 (C4), participants also did not have a session with robot but were asked to rest for 5 minutes.

A detailed experimental procedure for Study 2 is provided in section 4.8, page 121.

4.4.4 Effects of pet-robot therapy on heart rate and heart rate variability:

There are some studies of Paro that investigated its effects on heart rate (HR) and heart rate variability (HRV) (Mitsui et al. 2001*a*,*b*, Robinson et al. 2013c, Carretero & Sáez 2017).

As reviewed in section 2.3.3, page 24, in summary, past research findings found that when compared to baseline, interacting with an active Paro increased HRV (Mitsui et al. 2001*a*,*b*), but maintained levels of HR (Mitsui et al. 2001*a*, Carretero & Sáez 2017) or decreased the HR (Robinson et al. 2013c). It was claimed that the low HR was related to participants focusing on performing a task i.e. playing with active Paro (Mitsui et al. 2001*a*,*b*, Robinson et al. 2013c), and influenced by stroking Paro's soft fur (Robinson et al. 2013c). Then the high HRV in an active Paro condition was associated with positive experience as measured by self-reported moods (Mitsui et al. 2001*a*,*b*).

However, these past studies have not looked at the effects of interaction with Paro on physiological stress responses when preceded by or followed by a induced stress (i.e. noise exposure). But, based on the positive findings found in past studies of Paro, it was plausible that the Paro session might have an effect on reducing the effects of stress response. In Study 2, *it was anticipated that interacting with Paro might reduce the effects of induced stress based on improvement in psychological and physiological stress responses compare to stroking a bolster.*

4.5 Hypotheses

Based on the aims and rationale of Study 2 previously mentioned, the following are the hypotheses of this study (Refer to section 4.4.3, page 113 for summary of conditions in Study 2):

Hypothesis 1 (H1):

H1 aims to answer whether factors such familiarisation with Paro before intervention session may influenced participants' responses during intervention session, compared to other conditions.

It was predicted that in Session 1, a session with a Paro (C1 and C2) compared to a bolster (C3) or resting (C4) condition would result in:

H1(a) low self-reported stress, negative affect, negative emotions, and high self-reported positive affect and positive emotions (at at2).

H1(b) low HR and SCR, and high SDNN and HF (at t2).

H1(c) higher frequency of observed positive and lower negative facial expressions (at t2).

Hypothesis 2 (H2):
H2 aims to answer whether factors such familiarisation with Paro before intervention session may influenced participants' stress responses compared to other conditions, based on changes in scores between baseline and final-measurement (at3mat1 or t4mt1).

Based on changes in scores between baseline and final-measurement, it was predicted that a session with a Paro (C1 and C2) compared to bolster (C3) or resting (C4) condition would result in:

H2(a) a reduction in self-reported stress and negative affect, and an increase in self-reported positive affect. (at3mat1)

H2(b) a reduction in SCR, HR, and improvement in SDNN and HF. (t4mt1)

Then, two further hypotheses were made to find out if a short familiarisation session with Paro before the experiment (C1) will have a greater influence on participants' moods, physiological responses and ways that the participants interact with Paro during an intervention session, and its effect on subsequent stress exposure.

Hypothesis 3 (H3):
H3 was predicted based on scores in Session 1.

It was predicted that in Session 1, the effects of interacting with Paro in C1 compared to C2 would result in:

H3(a) low self-reported stress, negative affect, negative emotions, and high self-reported positive affect and positive emotions. (at at2)

H3(b) low HR and SCR, and high SDNN and HF. (at t2)

H3(c) higher frequency of observed positive and lower negative facial expressions. (at t2)

Hypothesis 4 (H4):
H4 was predicted based on changes in scores between baseline and final-measurement (at3mat1 or t4mt1).

Based on changes in scores between baseline and final-measurement, it was further predicted that effects of interacting with Paro (Session 1) in C1 compared to C2 would result in:

H4(a) a reduction in self-reported stress and negative affect, and an increase in self-reported positive affect. (at3mat1)

H4(b) a reduction in SCR, HR, and improvement in SDNN and HF. (t4mt1)

4.6 Methods

4.6.1 Participants

Ethical approval was obtained from University of Sheffield, Department of Computer Science Research Ethics Review Committees (Reference number: 011158). Experiments were conducted in an experiment room at Department of Psychology, University of Sheffield (See Figure 4.1 on page 118 for an overview of the experiment room). Participants were recruited from University of Sheffield.

Participants were recruited through University of Sheffield volunteering distribution-emailing list. A small announcement was also placed on Student Services information Desk (SSID) and On-line Research Participation system for Psychology students.

Participation was voluntary and all information was treated as confidential. Participants were told that they were free to leave at any time during experiment. Participants were offered the possibility of winning Amazon vouchers and Psychology students received 3 credits as a reward for their participation.

Exclusion criteria in Study 2 were similar to Study 1, as described in Chapter 3, section 3.7.1 on page 72 .

104 participants were recruited to take part in Study 2 (79 females, 25 males, participants were aged between 18 to 25, ($M = 19.24, SD = 1.782$). Table 4.1 shows summary of age and gender for each condition. A one-way between-subjects analysis of variance (ANOVA) test (gender as dependent variable and condition as between-subjects factor) found there was no significant difference in gender between conditions, $F(3,100) = 0.051, p = 0.985, r = 0.002$. A one-way between-subjects ANOVA test also (age as dependent variable and condition as between-subjects factor) found there was no significant difference in age between conditions, $F(3,100) = 0.416, p = 0.742, r = 0.012$.

Table 4.1: Report of age and gender by condition

Condition	*N*	Age	Gender
1	26	M = 19.46; SD = 1.923	M = 6; F = 20
2	26	M = 19.15; SD = 2.167	M = 6; F = 20
3	26	M = 18.96; SD = 1.248	M = 6; F = 20
4	26	M = 19.38; SD = 1.722	M = 7; F = 19
Total	104	M = 19.24; SD = 1.782	M = 25; F = 79

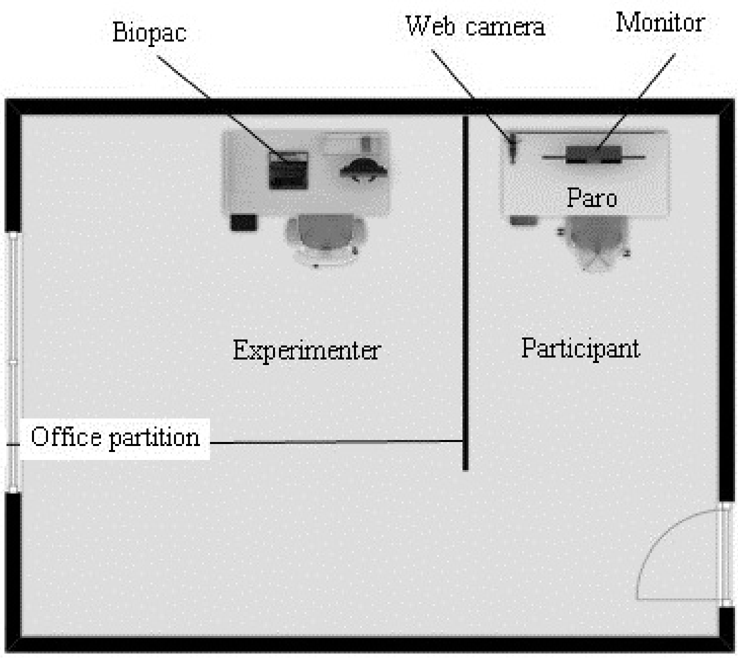


Figure 4.1: Experiment room at Department of Psychology. Notes: This is an overview of the experiment room and no exact measurement of the room was included. Participants and an experimenter were inside the same room, but were separated between an office partition. The measuring equipment (BIOPAC) used to collect the physiological data was on the table at the experimenter side. At the participant side, there was a monitor, a keyboard, a mouse and a web camera on the table, and the Paro or bolster was put on the table during the Session 1.

4.7 Measurement and materials

4.7.1 Bolster

A bolster with size dimension 14 x 48 centimetre. The size is quite similar to Paro, which is 35 x 16 x 57 centimetre (Paro Robots US 2014). The bolster has white faux soft fur.

Figure 4.2: The bolster used in condition 3 (C3)

4.7.2 Noise as a stressor

A new sound of babies crying was added to the combination of sounds used in Study 1. There were three noises used in this study: traffic noise, babies crying, crowd conversation.

The noises used in the experiment were downloaded from YouTube (www.youtube.com/watch?v=E_KPJhd857w, www.youtube.com/watch?v=mIXpLA4Q1Tc, https://www.youtube.com/watch?v=1UaJsxe3Lwg). The decibel levels of noise were edited using free software, Mp3Gain (http://mp3gain.sourceforge.net/).

The noises were combined into 5 minutes of sound. Then, to produce different levels of decibel in 5 minutes duration for each sound, the free software, Audacity were used (http://audacityteam.org/). The approximate levels of decibel (dB) were measured using the CR:160 Series Optimus Red sound level meter. The noises were delivered at an approximate volume of 60 to 80 dB through Sennheiser HD 202 headphones that connected to a laptop.

4.7.3 Physiological measurement and data acquisition

In Study 2, two physiological sensors were used to record participants' responses: skin conductance and pulse. A detailed description about skin conductance have been discussed earlier in this thesis, see Chapter 3, section 3.8.5 on page 75. Two simultaneous feedback channels were recorded using MP36, to record the real-time biophysical data acquisition processing. All sensors have a sampling rate of 200 Hz. Heart rate and heart rate variability data was acquired through photoplethysmography (PPG), see example of sensor in Figure 4.3. The sensor measures the blood volume using optical method by bouncing infra-red light against the surface of the skin to measure the amount of changes in light that reflected back. The pulses can be detected according to how much blood is in the skin, higher pulses resulted from more blood. In this study, the pulse sensor attached on the participants' thumb.

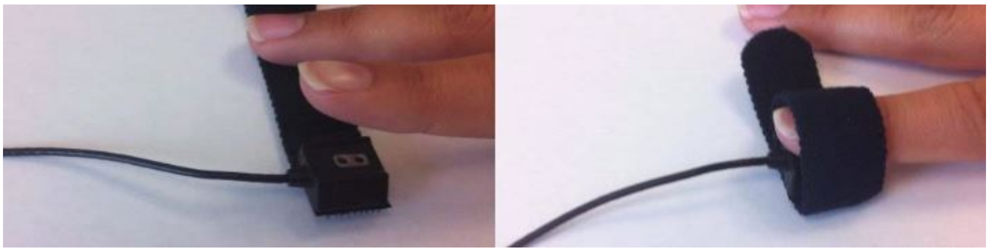


Figure 4.3: BIOPAC pulse sensor: Picture taken from BIOPAC Operator's Manual.

4.7.4 Video recording

Videos were recorded during the 5 minutes that participants were exposed to or interacted with the Paro and bolster. The videos were recorded using a web camera.

4.7.5 Reading material

Reading material a chapter from Music: A Very Short Introduction by Cook (2000) (See Appendix A.9) was given to participants in C2, C3, and C4. This is because participants in C1 were given a 2 minutes time to familiarise themselves with Paro. Participants in C2, C3 and C4 were required to spend 2 minutes reading and searching for the words "music" and "musician" in the reading material. The reading material was given to participants as a distraction task and to help them pass the 2 minutes time. This is also to avoid participants become disinterested with the task. Participants were told that the task is only for fun and they were not being graded.

4.7.6 Self-report of emotions

Self-reported ratings were used as a subjective measure of levels of emotions during Paro or bolster sessions. Participants were asked "*Please rate your emotions when you were playing with Paro or bolster.*" The question was asked after post-measurement (t4). Participants rated their feelings using 7-point Likert-scale ranging from "Not at all = 1" to "Extremely = 7" (Refer to Appendix A.14). This question was asked in condition 1 - 3 only.

4.7.7 Other materials

Other materials, including Paro, Positive and Negative Affect Schedule (PANAS) and levels of stress questionnaire, were identical to those used in Study 1. All detailed descriptions about Paro, PANAS and levels of stress have been discussed earlier in this thesis, see Chapter 3, section 3.8.1 on page 73, section 3.8.3 on page 74, and section 3.8.4 on page 74. However, in Study 2, question for levels of stress was rephrased to "Please indicate your stress levels right now."

4.8 Experimental procedure

Participants were asked to give their consent to participate in this study. Before the experiment starts, the experimenter explained how to play with the Paro robot. All instructions were displayed on the screen. Then participants were left in the room to complete a questionnaire with personal information e.g. age and gender. Afterwards participants were asked to sit comfortably and asked to wear headphones. Then participants were attached with physiological sensors on the fingertips (for heart rate variability (HRV) and skin conductance response (SCR)).

This study collected repeated measurements of skin conductance (SC) and HRV. Participants were also asked about their Positive and Negative Affect Schedule (PANAS) and levels of stress. The HRV and SC were measured for 5 minutes at four time points. A pre-measurement (t1) was followed by Session 1 (Familiar Paro, Paro, bolster, rest) (t2), a 5 minutes of Session 2 (noise induction session) (t3) and a post-measurement (t4). Participants' session with robot or bolster were recorded on video. Then the PANAS questionnaires and levels of stress questionnaires were collected after pre-measurement (at1), after Session 1 (at2) and after Session 2 (at3).

Participants were assigned randomly to one of four conditions. There are four different conditions in Session 1 in this study.

In condition 1 (C1), before the experiment starts, the experimenter left the participant in the room and gave them 2 minutes to play with Paro robot before the experimenter started collecting their responses (t1). Then again they were given a time with the robot for 5 minutes during the Session 1 (t2). For condition 2, 3 and 4, participants were instead given some reading material, and asked to find and circle the words "music" and "musician" in the reading material. In condition 2 (C2), they were given a 5 minutes to play with the robot during the robot session. In condition 3 (C3), participants will not have a session with the robot but participants were asked to stroke the furry bolster. The bolster is covered with soft fur which quite similar to soft texture of the Paro robot. In condition 4 (C4), participants did not have a session with the robot but were asked to sit comfortably for 5 minutes. This final condition was assigned as control condition, in order to check if the noise induced mild stress.

The four conditions were labelled as follow: C1:familiar, C2:no_familiar, C3:bolster, C4:rest.

During the noise session, participants were asked to listen to the sound through headphones for 5 minutes. Participants were allowed to ask the experimenter to stop the experiment if they found the sound too stressful. However, none of the participants pleaded to stop the experiment early.

After completing all the sessions, participants were asked to answer final questionnaire. After they finished answering all questionnaires, all the sensors attached to participants were removed. Before leaving the room, participants were given time to ask further questions about this study. Figure 4.4 shows the flow of the experiment.

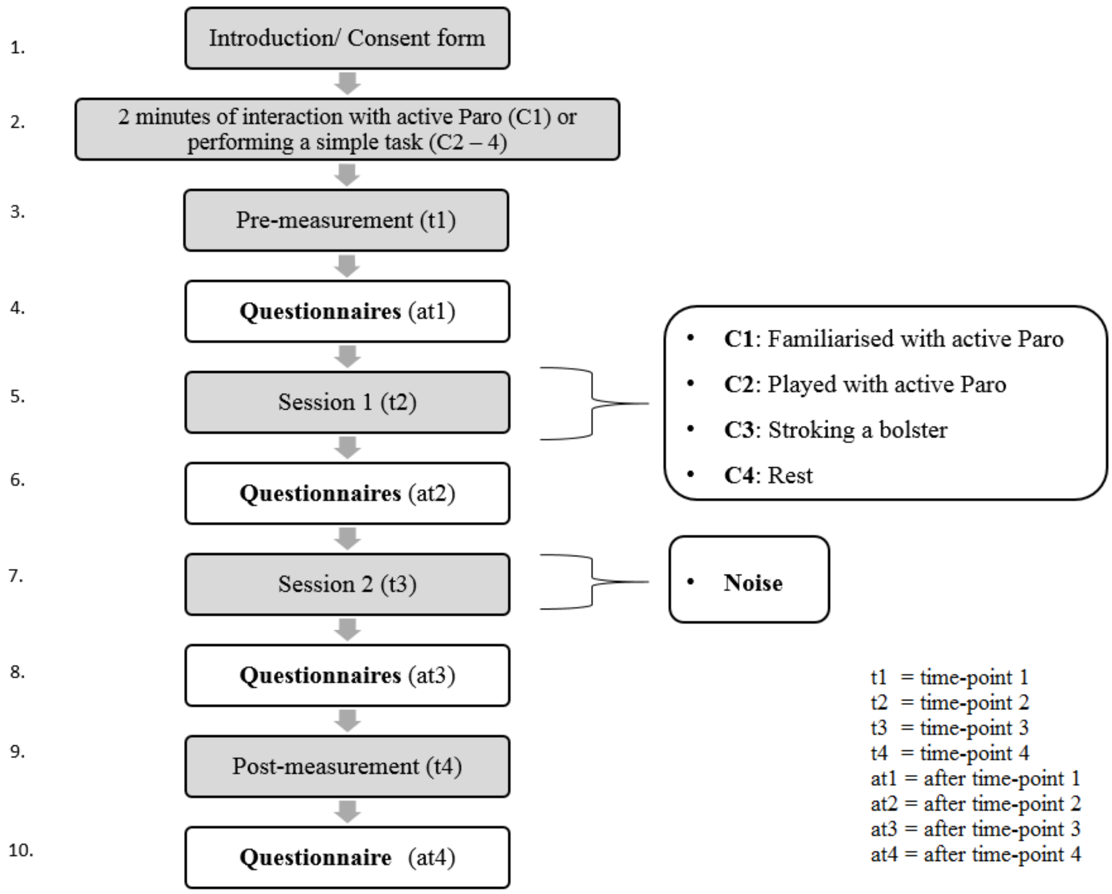


Figure 4.4: Study 2 experimental procedure

4.9 Data analysis

Skin conductance data:

Skin conductance data were calculated by totalling up the frequency of peaks during the 5 minutes recordings for each time point (pre-measurement (t1), Session 1 (t2), Session 2 (t3), post-measurement (t4)). Skin conductance data were analysed using a similar way in Study 1, refer section 3.10 on page 78.

Heart rate and HRV data pre-processing:

Then for photoplethysmography (PPG) data, the heart rate scores were pre-processed using Kubios 2.0 software (Tarvainen et al. 2014). Kubios is a free software that downloaded from https://www.kubios.com/. The signals were corrected using the default option "medium artefact correction" in Kubios software. Additionally, to remove disturbing low frequency baseline trend components, a Lambda value of 500 was set to the "smoothness prior method" option (Luque-Casado et al. 2013). Then the data were manually inspected and corrected for any artefact on or removal (Krypotos et al. 2011).

Then the data from heart rate variability were pre-processing using AcqKnowledge 4.4 software. The data noises from PPG signal (i.e. movement) were removed using high-pass filters with cut-off frequencies of 0.5 Hz (Lu et al. 2008, Jeyhani et al. 2015).

Data normality check:

Data analysis was done separately for physiological and psychological measures. All data were checked for normality using Shapiro-Wilk test (Razali et al. 2011). The data were carefully inspected using box plots. Extreme outliers were removed (scores that were more than three standard deviation away from the mean (Howitt 2017).

Baseline:

First, the data were analysed in order to check that was no difference between the participants in the four conditions (C1:familiar, C2:no_familiar, C3:bolster, C4:rest) at the pre-measurement (t1) for all variables (skin conductance responses, PANAS, stress levels). A one-way between-subjects ANOVA was performed with pre-measurement (taken at t1 for physiological responses and at at1 for self-reported responses) as dependent variable and conditions as a between-subjects factor. The alpha level was set at $p < 0.05$.

Overall effect of time and conditions:

Then the data were analysed to look at the effects of time and condition on the physiological and self-reported responses. A mixed between-within subjects ANOVA was performed for each measure with four conditions (C1:familiar, C2:no_familiar, C3:bolster, C4:rest) as a between-subjects factor and four time points (t1, t2, t3, t4) as within-subjects variable factor. The dependent variables were self-reported stress, positive affect, negative affect and skin conductance. For each analysis, any ANOVA variables found to violate the assumption of sphericity were corrected with a Greenhouse-Geisser correction and all post-hoc tests were done with a Bonferroni correction. The alpha level was set at $p < 0.05$.

Then in order to look at a response pattern of the variables over the time period for mixed design between-within subjects ANOVA, a polynomial contrast test was used. For example, time pattern from t1 to t2, from t2 to t3, and from t3 to t4. Additionally, looking at time differences between baseline or pre-measurement, i.e. from t1 to t4.

New variables:

For further analysis, in order to assess if the changes of scores between pairs of time points are differ across four different conditions, new variables were created based on pairs of time points. A one-way between-subjects analysis of variance was performed with each new variable as dependent variable and conditions as a between-subjects factor. A further explanation is reported on each subsection of the results.

Self-reported emotions:

Next analysis for self-reported were done according to the paragraph above, but additional analysis was done. There were 6 emotional states that participants were asked about: 3 negative ones (boredom, disgust, confusion) and 3 positive ones (happiness, excitement, relaxation). New variables were created in order to compare the results from positive and negative emotions between conditions, the analysis was adapted from a study by (Takayanagi et al. 2014). The scores from happy, excitement, relaxation were totalled up and the variable was named as positive emotions. The scores from boredom, disgust, confusion were totalled up and the variable was named as negative emotions.

Observed facial emotional expressions:

Then analysis for observed facial emotional expression. Observation was undertaken 5 minutes of video recorded during the session with Paro and bolster in C1:familiar, C2:no_familiar and C3:bolster. There was no video recording during the rest session in C4:rest. There were four facial expressions that were checked and divided into positive emotion (*smile and laugh*) and negative emotion (*boredom and disgust*). The 5 minutes of videos were divided into 30 blocks of 10 seconds, and were analysed by two evaluators for inter rater reliability. The inter-rater reliability for all emotions were analysed using Intraclass Correlation Coefficient (ICC) (Landers 2015). Then after checking the inter-rater reliability, the coded facial expression scores were totalled for each emotion for each condition, using the mean scores of two evaluators for further analysis (Takayanagi et al. 2014). Other analysis for observed video were done similarly to Study 1, refer section 3.10 on page 78.

4.10 Results

4.10.1 Baseline of each measure

There was no significant difference between conditions at pre-measurement (at1) in all the self-reported and physiological responses. These results indicate that participants in all conditions had a similar level of self-reported stress, Positive and Negative Affect Schedule (PANAS) and physiological responses at the beginning of the study, all cases $p > 0.05$.

4.10.2 Effects of conditions on self-reported responses

Self-reported stress

A mixed between-within subjects analysis of variance showed there was a significant effect of time on stress levels, $F(1.710, 171.00) = 78.09, p < 0.0001, r = 0.439$, and no significant effect of conditions, $F(3, 100) = 2.006, p = 0.118, r = 0.057$. There was no significant interaction between condition and time on stress levels, $F(5.130, 171.00) = 1.481, p = 0.197, r = 0.043$.

Pairwise comparisons between time points were done using Bonferroni correction test. Based on Figure 4.5, the stress level scores showed a significant decrease during Session 1 (at2) from pre-measurement (at1), $p < 0.0001$, a significant increase during the Session 2 (at3) from the Session 1 (at2), $p < 0.0001$, and a significant increase between pre-measurement (at1) and Session 2 (at3), $p = 0.040$.

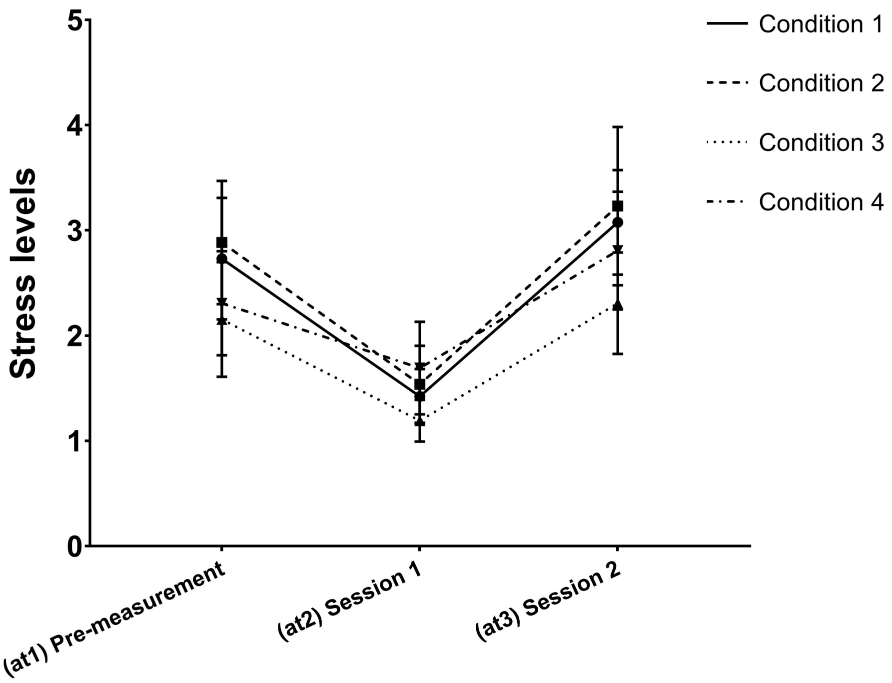


Figure 4.5: The mean with 95% confidence interval for self-reported ratings of levels of stress to each condition.

New variables: New variables were created in order to check if there is difference between conditions in terms of changes in stress level scores. The new variables are $SL\_at2mat1$ = Session 1 minus pre-measurement, $SL\_at3mat2$ = Session 2 minus Session 1 and $SL\_at3mat1$ = Session 2 minus pre-measurement.

There was a significant difference between conditions for $SL\_at2mat1$, $F(3,100) = 3.511, p = 0.018, r = 0.095$. It was found that the decrease in stress level scores were significantly larger in C2:no_familiar compared to C4:rest , $p = 0.034$. (Refer to Figure 4.6). There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, C1:familiar and C3:bolster, $p = 1.000$, C1:familiar and C4:rest, $p = 0.051$, C2:no_familiar and C3:bolster, $p = 0.835$ and C3:bolster and C4:rest, $p = 1.000$.

There was no significant difference between conditions for $SL\_at3mat2$, $F(3,100) = 1.793, p = 0.153, r = 0.051$. the stress level scores increase from Session 1 (at2) to the Session 2 (at3) for all conditions. (Refer to Figure 4.6).

There was no significant difference between conditions for $SL\_at3mat1$, $F(3,100) = 0.281, p = 0.839, r = 0.008$. There no changes in stress level scores were found between pre-measurement (at1) and Session 2 (at3) for all conditions.

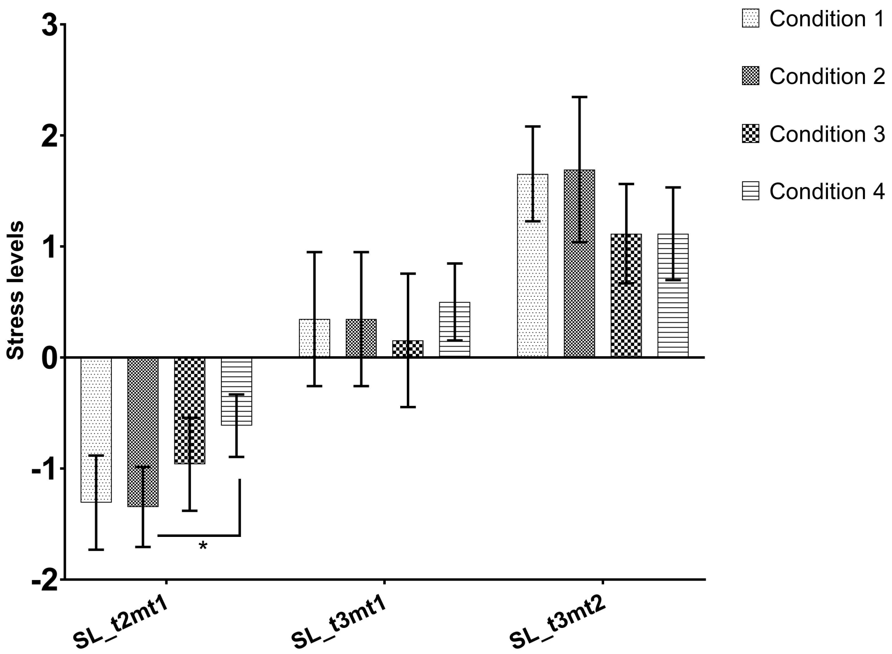


Figure 4.6: The mean with 95% confidence interval for changes of scores of levels of stress to each condition. $*p < 0.05$

Session 1 (at2) and Session 2 (at3) There was no significant difference between conditions at Session 1, $F(3,100) = 1.738, p = 0.164, r = 0.050$.

There was no significant difference between conditions at Session 2, $F(3,100) = 2.051, p = 0.112, r = 0.058$. Although there was no significant difference as can be seen in Figure 4.11, page 136, in Session 2, the stress levels in C3:bolster were lower compared to other conditions.

Positive affect in PANAS

A mixed between-within subjects ANOVA showed there was a significant effect of time on positive affect, $F(2,200) = 80.03, p < 0.0001, r = 0.445$, and a significant effect of conditions, $F(3,100) = 2.687, p = 0.040, r = 0.079$. There was a significant interaction between condition and time on positive affect, $F(6,200) = 8.831, p < 0.0001, r = 0.209$.

Pairwise comparisons between time points were done using the Bonferroni correction test. The was no significant difference in positive affect scores after Session 1 (at2) compared to the pre-measurement (at1), $p = 0.561$. The positive affect scores then showed a significant decrease after the Session 2 (at3) from the Session 1 (at2), $p < 0.0001$. There was a significant decrease in the positive affect scores between pre-measurement (at1) and Session 2 (at3), $p < 0.0001$. Refer to Figure 4.7.

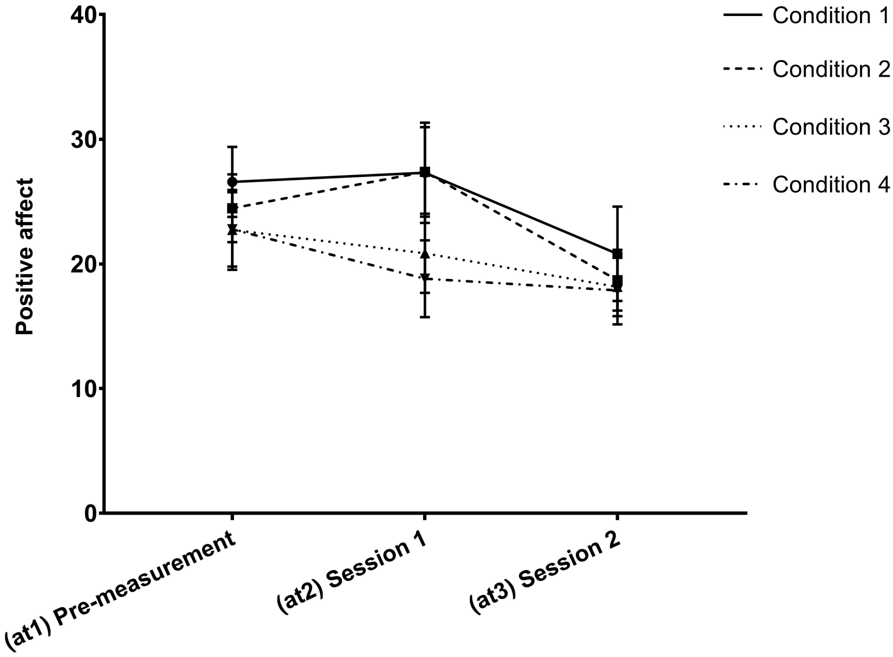


Figure 4.7: The mean with 95% confidence interval for positive affect in PANAS for each condition.

New variables: New variables were created in order to check if there is a difference between conditions in terms of changes in positive affect scores. The new variables are $PA\_at2mat1$ = Session 1 minus pre-measurement, $PA\_at3mat2$ = Session 2 minus Session 1 and $PA\_at3mat1$ = Session 2 minus pre-measurement.

There was a significant difference between conditions for $PA\_at2mat1$, $F(3,100) = 13.28, p < 0.0001, r = 0.285$. The increase in positive affect scores were significantly higher in C1:familiar compared to C4:rest, $p < 0.001$, in C2:no_familiar compared to C3:bolster,

$p < 0.0001$, and in C2:no_familiar compared to C4:rest, $p < 0.001$. However, C3:bolster and C4:rest showed a decrease in positive affect scores. (Refer to Figure 4.8).

However, there were no significant differences between C1:familiar and C2:no_familiar, $p = 0.379$, C1:familiar and C3:bolster, $p = 0.163$, and C3:bolster and C4:rest, $p = 0.469$.

There was also a significant difference between conditions for $PA\_at3mat2$, $F(3,100) = 12.77, p < 0.0001, r = 0.277$. The decrease in positive affect scores was significantly larger in C1:familiar compared to C3:bolster, $p = 0.046$, and to C4:rest, $p = 0.001$, and C2:no_familiar compared to C3:bolster, $p < 0.001$, and C2:no_familiar compared to C4:rest, $p < 0.001$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.721$, and C3:bolster and C4:rest, $p = 1.000$. Although it was not significant, as can be seen in Figure 4.8, C3:bolster and C4:rest showed a further decrease in positive affect,

There was also no significant difference between conditions for $PA\_at3mat1$, $F(3,100) = 0.450, p = 0.718, r = 0.013$. As seen in Figure 4.8, all conditions showed a further decrease in positive affect scores after the Session 2 (at3) as compared to the pre-measurement session (at1).

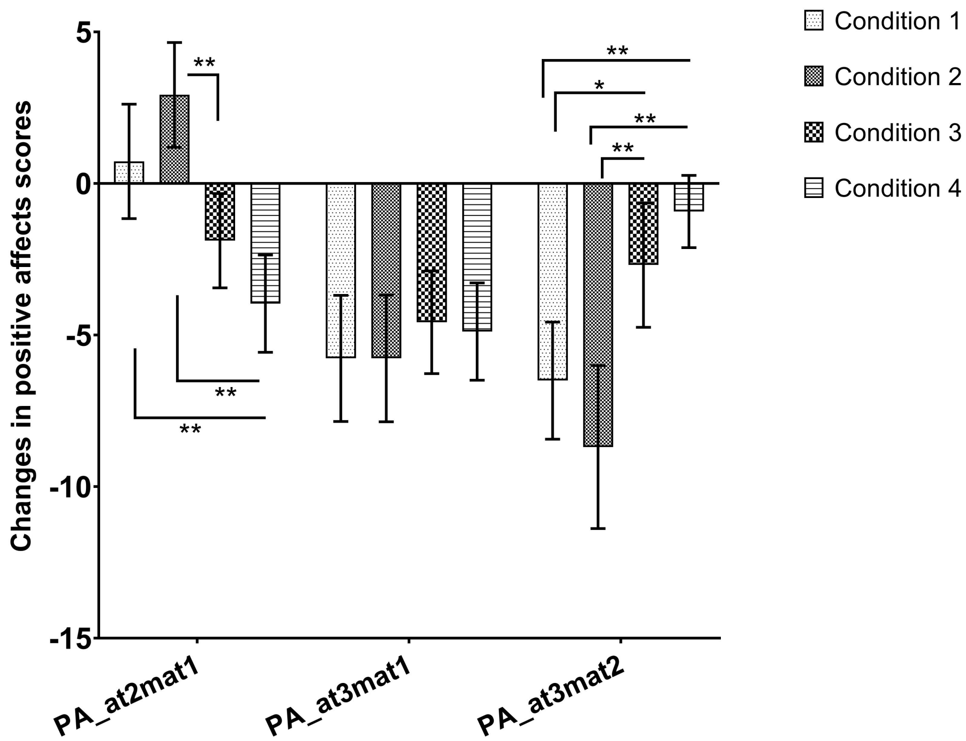


Figure 4.8: The mean with 95% confidence interval for changes in positive affect scores in PANAS for each condition. $*p < 0.05, **p < 0.001$

Session 1 (at2) and Session 2 (at3) There was a significant difference between conditions at Session 1, $F(3,100) = 6.827, p < 0.0001, r = 0.170$. The positive affect was significantly higher in C1:familiar compared to C3:bolster, $p = 0.049$ and C4:rest, $p = 0.004$. The results were similar for C2:no_familiar, the positive affect was significantly higher in C2:no_familiar compared to C3:bolster, $p = 0.045$ and C4:rest, $p = 0.003$. (Refer to Figure 4.11 on page 136). There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, and C3:bolster and C4:rest, $p = 1.000$.

There was no significant difference between conditions at Session 2, $F(3,100) = 0.895, p = 0.446, r = 0.026$. Although there was no significant difference, as can be seen in Figure 4.11 on page 136, the PA in C1:familiar was higher compared to other conditions.

Negative affect in PANAS

Eight outliers were removed from data analysis.

A mixed between-within subjects ANOVA showed there was a significant effect of time on negative affect, $F(1.448, 136.09) = 63.36, p < 0.0001, r = 0.403$, and no significant effect of conditions, $F(3, 94) = 1.429, p = 0.239, r = 0.044$. There was no significant interaction between condition and time on negative affect, $F(4.343, 136.09) = 1.364, p = 0.231, r = 0.042$.

Pairwise comparisons between time points were done using Bonferroni correction test. The negative affect scores decrease significantly from pre-measurement (at1) to Session 1 (at2), $p < 0.0001$. The negative affect scores showed significant increase from Session 1 (at2) to Session 2 (at3), $p < 0.0001$. The negative affect scores also showed significant increase from pre-measurement (at1) to Session 2 (at3), $p < 0.0001$.

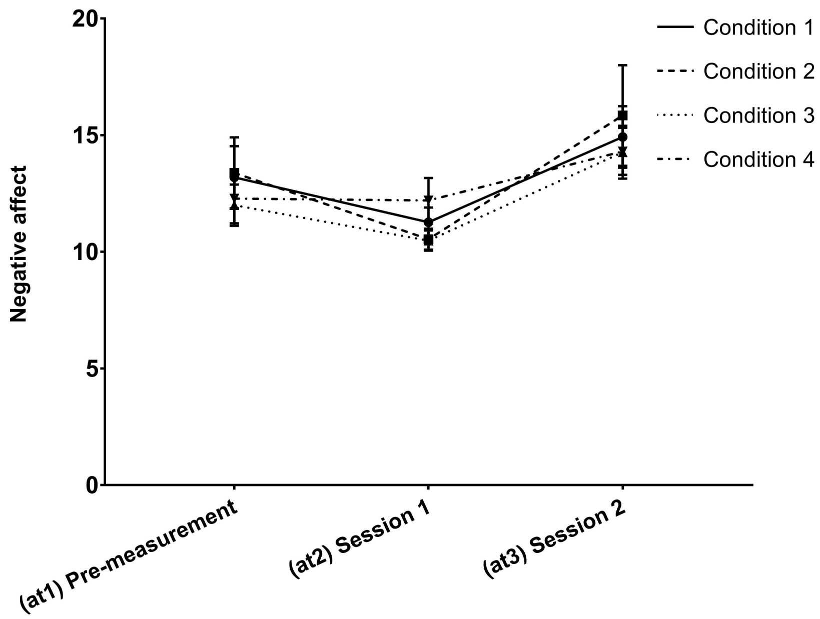


Figure 4.9: The mean with 95% confidence interval for negative affect in PANAS measures to each condition.

New variables: New variables were created in order to check if there is difference between conditions in terms of changes in negative affect scores. The new variables are $NA\_at2mat1$ = Session 1 minus pre-measurement, $NA\_at3mat1$ = Session 2 minus pre-measurement and $NA\_at3mat2$ = Session 2 minus Session 1.

There was a significant difference between conditions for $NA\_at2mat1$, $F(3,96) = 5.149, p = 0.002, r = 0.139$. The negative affect scores showed a significantly larger decrease from pre-measurement (at1) to the noise session (at2) in C1:familiar and compared to C4:rest, $p = 0.016$, and C2:no_familiar compared to C4:rest, $p = 0.003$. (Refer to Figure 4.10). However, there were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, C1:familiar and C3:bolster, $p = 1.000$, C2:no_familiar and C3:bolster, $p = 0.527$ and C3:bolster and C4:rest, $p = 0.346$.

There was no significant difference between conditions for $NA\_at3mat2$, $F(3,94) = 1.739, p = 0.164, r = 0.053$. As evident in Figure 4.10, the negative affect scores increase from Session 1 (at2) to Session 2 (at3) for all conditions.

There was no significant difference between conditions for $NA\_at3mat1$, $F(3,97) = 0.179, p = 0.910, r = 0.006$. As shown in Figure 4.10, the negative affect scores increase from pre-measurement (at1) to Session 2 (at3) for all conditions.

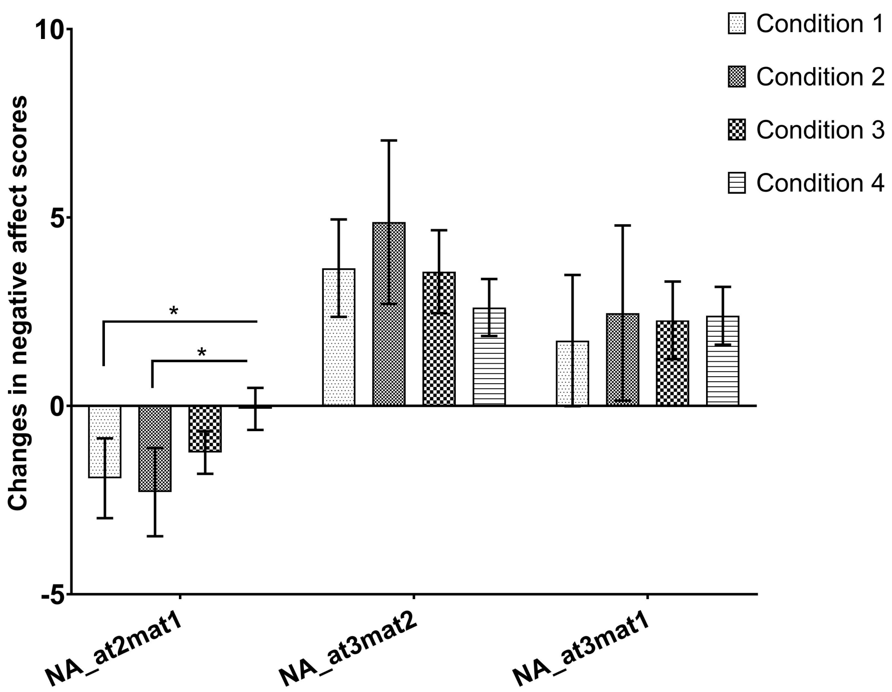


Figure 4.10: The mean with 95% confidence interval for changes in negative affect scores in PANAS for each condition. $*p < 0.05$

Session 1 (at2) and Session 2 (at3) There was a significant difference between conditions at Session 1, $F(3,96) = 6.261, p = 0.001, r = 0.14$. The negative affect was higher in C4:rest compared to C2:no_familiar, $p = 0.003$ and C4:rest compared to C3:bolster, $p = 0.004$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.667$, C1:familiar and C3:bolster, $p = 0.487$, C1:familiar and C4:rest, $p = 0.242$, and C2:no_familiar and C3:bolster, $p = 1.000$.

There was no significant difference between conditions at Session 2, $F(3,97) = 1.039, p = 0.379, r = 0.031$. Although there was no significant difference, as can be seen in Figure 4.11, the negative affect in C1:familiar and C2:no_familiar was higher compared to other conditions.

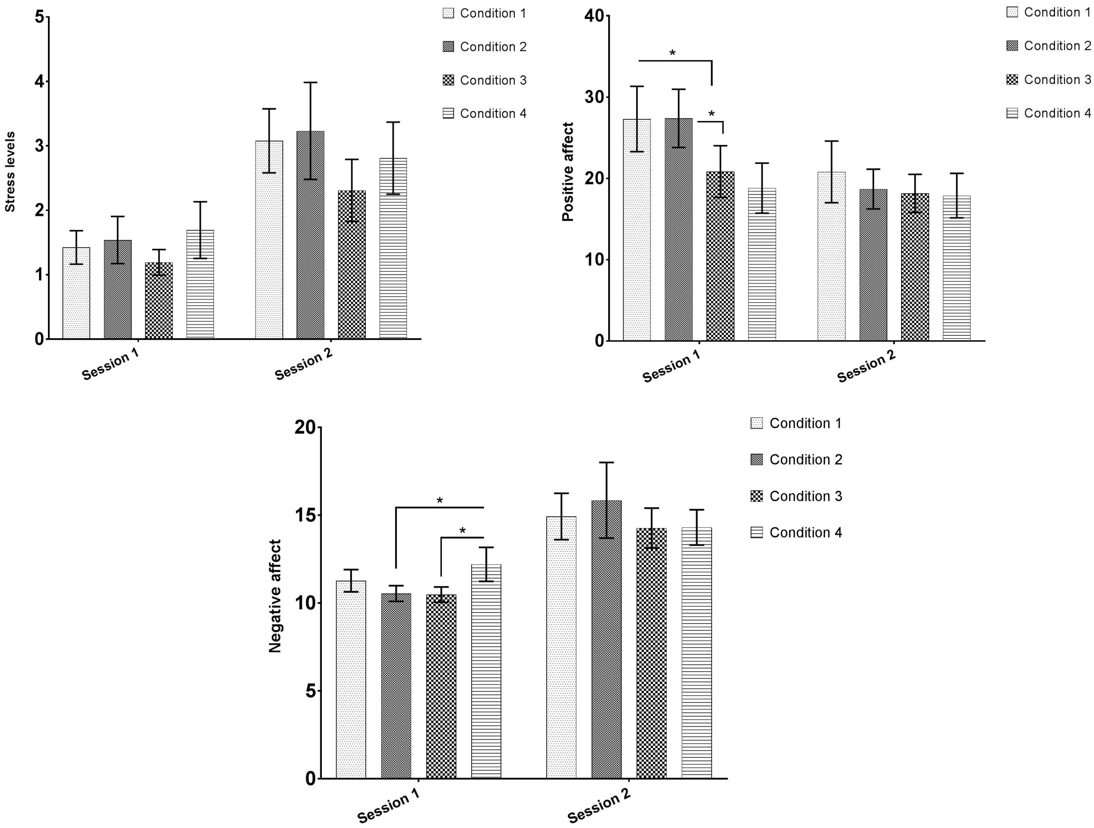


Figure 4.11: The mean with 95% confidence interval for Session 1 and Session 2 of self-reported stress, positive and negative affect to each condition. $*p < 0.05$.

Self-reported emotions during Session 1

There were 6 emotional states that participants were asked about: 3 negative ones (boredom, disgust, confusion) and 3 positive ones (happiness, excitement, relaxation). A one-way between-subjects analysis of variance was used to analyse difference between conditions for each emotional state. There were no significant differences between conditions on self-reported responses of happy, relaxed, bored, and disgusted during Session 1 (at2), all cases $p > 0.05$.

There was a significant difference between conditions on the self-reported emotion excited during Session 1 (at2), $F(2,75) = 13.37, p < 0.0001, r = 0.269$. A post hoc test using Bonferroni correction found a significant difference between C1:familiar and C3:bolster, $p < 0.0001$

and between C2:no_familiar and C3:bolster, $p < 0.0001$. The excited scores were higher in C1:familiar and C2:no_familiar compared to C3:bolster. (Refer to Figure 4.12). There was no significant difference between C1:familiar and C3:bolster, $p = 1.000$.

A significant difference between conditions also found on self-reported emotion confused during Session 1 (at2), $F(2,75) = 4.841, p = 0.011, r = 0.114$. A post hoc using Bonferroni test found a significant difference between C1:familiar and C2:no_familiar, $p = 0.019$, and between C1:familiar and C3:bolster, $p = 0.036$. The confused scores were lower in C1:familiar compared to C2:no_familiar and C3:bolster. (Refer to Figure 4.12). There was no significant difference between C2:no_familiar and C3:bolster, $p = 1.000$.

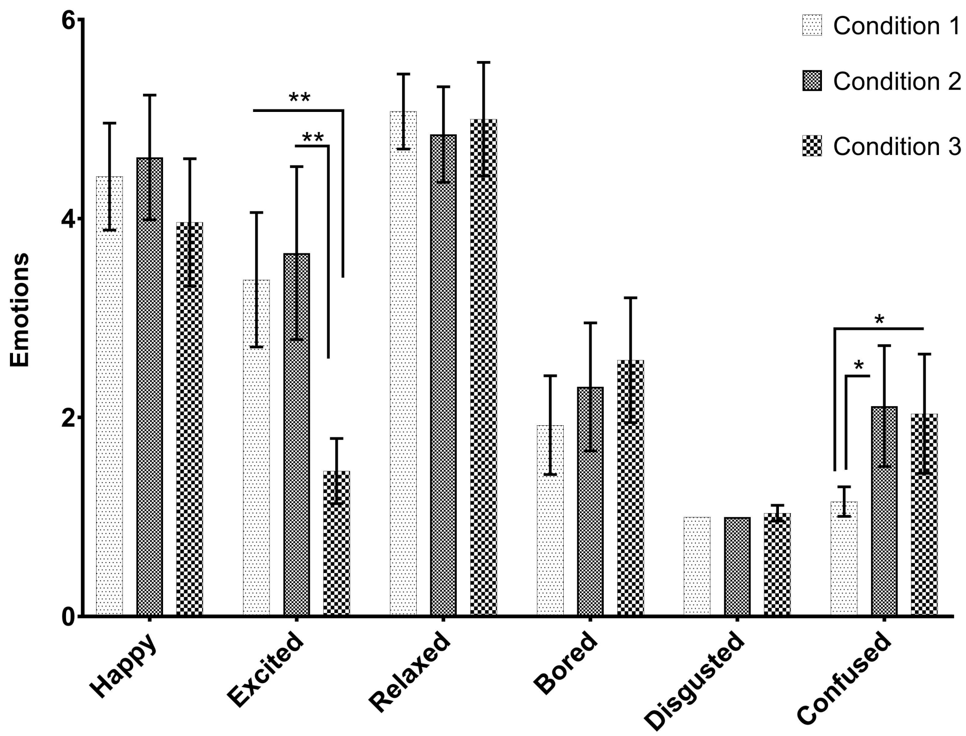


Figure 4.12: The mean with 95% confidence interval for self-reported emotions during Session 1 to each condition. $*p < 0.05, **p < 0.001$

A further analysis was done to assess effects of conditions for positive and negative emotions.

There was a significant difference between conditions on positive emotion during Session

1 (at2), $F(2,75) = 4.621, p = 0.013, r = 0.110$. A post hoc test using Bonferroni correction found a significant difference between C1:familiar and C3:bolster, $p < 0.0001$ and between C2:no_familiar and C3:bolster, $p < 0.0001$. The positive emotions scores were higher in C1:familiar and C2:no_familiar compared to C3:bolster. (Refer to Figure 4.13). There was no significant difference between C1:familiar and C3:bolster, $p = 1.000$.

There was a significant difference between conditions on negative emotion during Session 1 (at2), $F(2,75) = 4.652, p = 0.012, r = 0.110$. A post hoc test using Bonferroni correction found a significant difference between C1:familiar and C3:bolster, $p < 0.0001$. The negative emotions scores were lower in C1:familiar compared to C3:bolster. (Refer to Figure 4.13). There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.055$, and between C2:no_familiar and C3:bolster, $p = 1.000$.

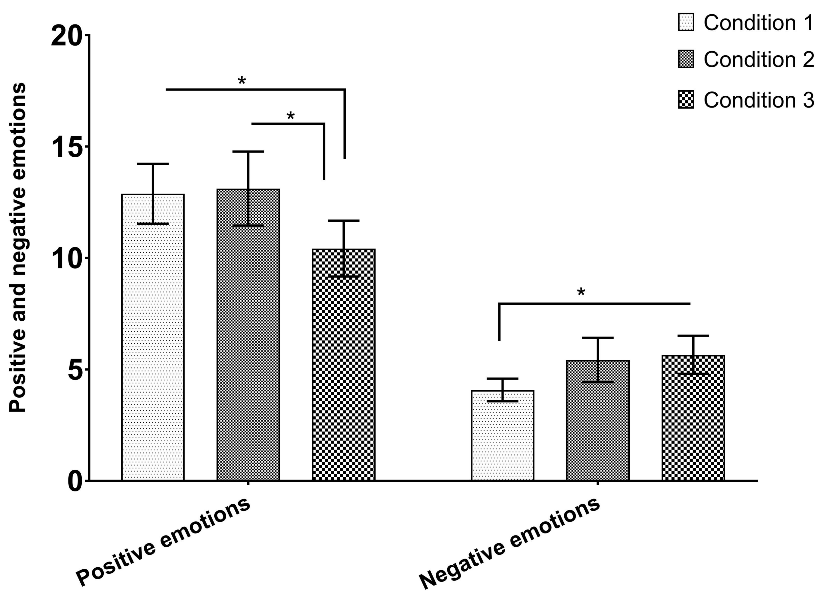


Figure 4.13: The mean with 95% confidence interval for positive and negative expression on self-reported emotions during Session 1 to each condition. $*p < 0.05$

4.10.3 Effects of conditions on observed facial emotional expressions

Two outliers were removed, one outlier from C1:familiar and C4:rest. A moderate degree of reliability was found between two evaluators for the coded facial expressions, α between 0.724 to 0.830. However, no analysis for the coded facial expression, disgust, because there was not enough data for this analysis.

A one-way between-subjects ANOVA test was used to test for differences between conditions for each coded expression: smile, laugh, disgust and boredom. There was a significant difference between conditions in the number of times each of the following expressions occurred: smile $F(2,72) = 29.49, p < 0.0001, r = 0.450$, laugh $F(2,72) = 4.689, p = 0.012, r = 0.115$ and boredom $F(2,72) = 22.09, p < 0.0001, r = 0.380$. There was no significant difference between conditions for the coded facial expression, disgust $F(2,72) = 0.711, p = 0.494, r = 0.019$.

Further post hoc analysis with Bonferroni correction was done to significant results above. For the coded facial expression, smile, a post-hoc comparison found a significant difference between C1:familiar and C3:bolster, $p < 0.0001$ and between C2:no_familiar and C3:bolster, $p < 0.0001$. The smiling scores were higher in C1:familiar compared to C3:bolster. The smiling scores also higher in C2:no_familiar compared to C3:bolster. (Refer to Figure 3.10). There was no significant difference between C1:familiar and C2:no_familiar, $p = 1.000$.

For the coded facial expression, laugh, a post-hoc comparison found a significant difference between C2:no_familiar and C3:bolster, $p = 0.013$. The laughing scores were higher in C2:no_familiar compared to C3:bolster. (Refer to Figure 3.10). There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$ and C1:familiar and C3:bolster, $p = 0.100$.

For the coded facial expression, boredom, a post-hoc comparison found a significant difference between C1:familiar and C3:bolster, $p < 0.0001$ and between C2:no_familiar and C3:bolster, $p < 0.0001$. The boredom scores were lower in C1:familiar compared to C3:bolster.

The boredom scores were also lower in C2:no_familiar compared to C3:bolster. (Refer to Figure 3.10). There was no significant difference between C1:familiar and C2:no_familiar, $p = 1.000$.

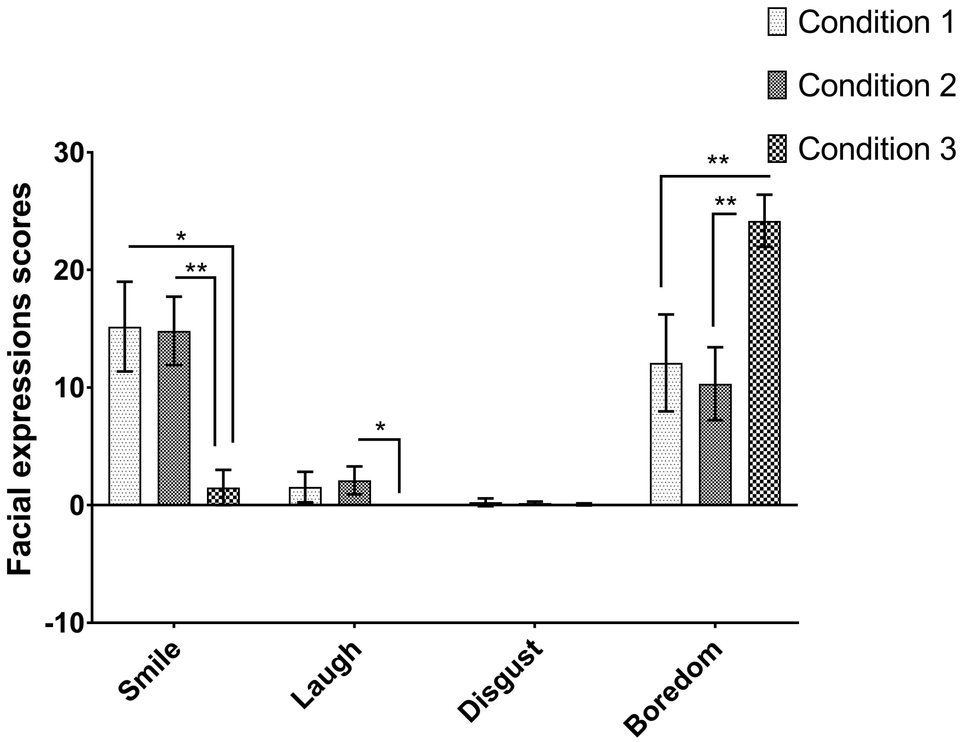


Figure 4.14: The mean with 95% confidence interval for facial emotional expressions to each condition. $*p < 0.05, **p < 0.001$

A further analysis was done to compare scores for positive and negative facial expressions for each condition.

There was a significant difference between conditions for positive facial expressions, $F(2,72) = 31.85, p < 0.0001, r = 0.469$. A post hoc test was done using Bonferroni correction. As shown in Figure 4.15, there was a significant difference between C1:familiar and C3:bolster, $p < 0.0001$ and between C2:no_familiar and C3:bolster, $p < 0.0001$. The positive expressions scores were higher in C1:familiar compared to C3:bolster and in C2:no_familiar compared to C3:bolster. There was no significant difference between C1:familiar and C2:no_familiar, $p = 1.000$.

There was a significant difference between conditions for negative facial expressions, $F(2,72) = 21.72, p < 0.0001, r = 0.376$. A post hoc test was done using Bonferroni correction. As shown in Figure 4.15, there was a significant difference between C1:familiar and C3:bolster, $p < 0.0001$ and between C2:no_familiar and C3:bolster, $p < 0.0001$. The negative expressions scores were lower in C1:familiar compared to C3:bolster and in C2:no_familiar compared to C3:bolster. There was no significant difference between C1:familiar and C2:no_familiar, $p = 1.000$.

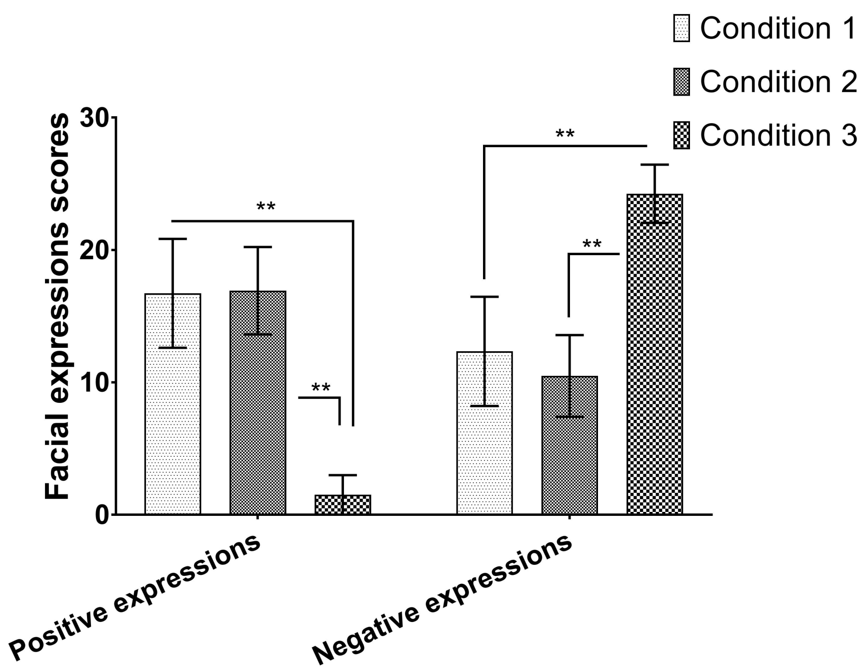


Figure 4.15: The mean with 95% confidence interval for negative and positive expressions to each condition. $*p < 0.05, **p < 0.001$

4.10.4 Effects of conditions on physiological responses

Skin conductance responses

A mixed between-within subjects analysis of variance showed there was a significant effect of time on SCR, $F(2.591, 259.11) = 214.43, p < 0.0001, r = 0.682$, and a significant effect of conditions, $F(3,100) = 10.94, p < 0.0001, r = 0.247$. There was significant interaction between condition and time on SCR, $F(7.773, 259.11) = 29.59, p < 0.0001, r = 0.470$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. There was a significant increase in SCR from pre-measurement (t1) to Session 1 (t2), $p < 0.0001$. The SCR showed a significant decrease from Session 1 (t2) to Session 2 (t3), $p < 0.0001$. The SCR showed no significant difference in SCR from pre-measurement (t1) to post-measurement (t4), $p = 1.000$ and from Session 2 (t3) to post-measurement (t4), $p = 0.468$.

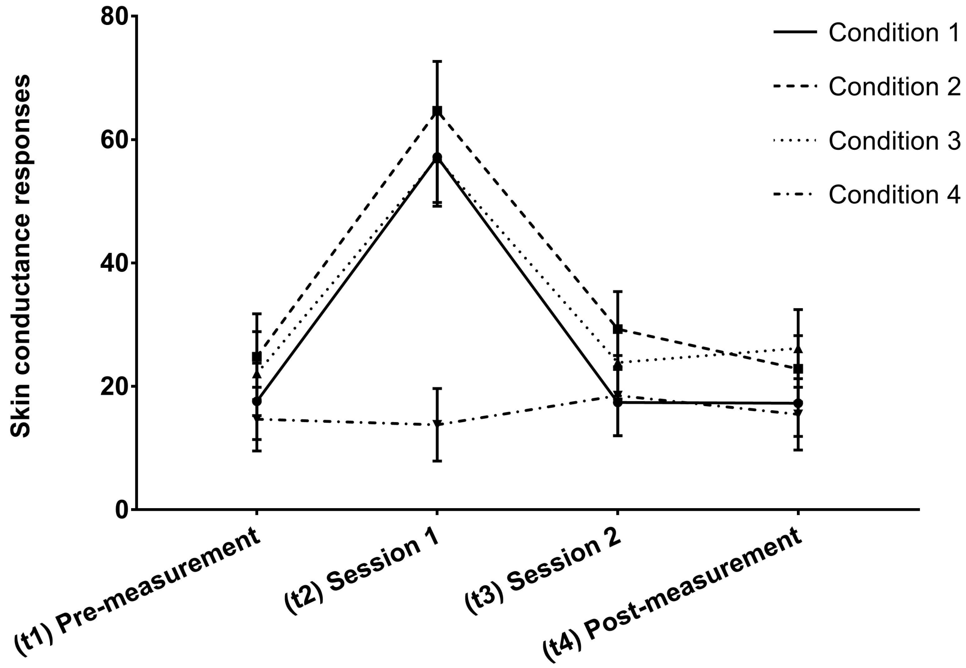


Figure 4.16: The mean with 95% confidence interval for skin conductance responses to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in SCR. The new variables are SCR_t2mt1 = Session 1 minus pre-measurement, SCR_t3mt2 = Session 2 minus Session 1, SCR_t4mt3 = post-measurement minus Session 2, and SCR_t4mt1 = post-measurement minus pre-measurement.

There was a significant difference between conditions for SCR_t2mt1, $F(3, 100) =$

$38.33, p < 0.0001, r = 0.535$. As is apparent in Figure 4.17, the SCR increased in C1:familiar, C2:no_familiar and C3:bolster compared to C4:rest, all cases with $p < 0.0001$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, C1:familiar and C3:bolster, $p = 1.000$, and C2:no_familiar and C3:bolster, $p = 1.000$.

There was a significant difference between conditions for SCR_t3mt2, $F(3,100) = 52.63, p < 0.0001, r = 0.612$. As evident in Figure 4.17, the decrease in SCR was a significantly larger in C1:familiar, C2:no_familiar and C3:bolster compared to C4:rest, all cases with $p < 0.0001$. C4:rest showed a small increase in SCR. There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, C1:familiar and C3:bolster, $p = 0.593$, and C2:no_familiar and C3:bolster, $p = 1.000$.

There was a significant difference between conditions for SCR_t4mt3, $F(3,100) = 3.408, p = 0.021, r = 0.093$. As shown in Figure 4.17, there was a small decrease SCR in C2:no_familiar compared to C3:bolster, $p = 0.019$. The C3:bolster showed a small increase in SCR. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.188$, C1:familiar and C3:bolster, $p = 1.000$, C1:familiar and C4:rest, $p = 1.000$, C2:no_familiar and C4:rest, $p = 1.000$ and C3:bolster and C4:rest, $p = 0.402$.

There was no significant difference between conditions for SCR_t4mt1, $F(3,100) = 1.097, p = 0.354, r = 0.032$.

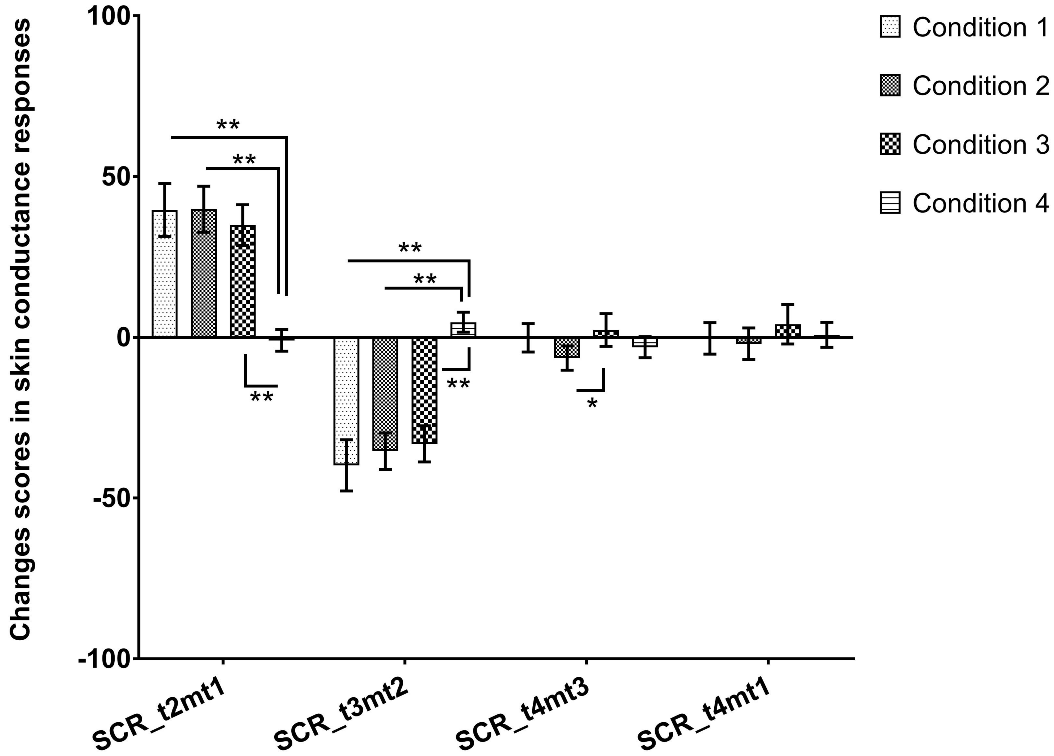


Figure 4.17: The mean with 95% confidence interval for changes in scores of skin conductance responses to each condition. $*p < 0.05, **p < 0.001$

Session 1 (t2) and Session 2 (t3) There was a significant difference between conditions at Session 1, $F(3,100) = 42.73, p < 0.0001, r = 0.562$. As can be seen in Figure 4.24 on page 154, the SCR for C4:rest was lower compared to C1:familiar, C2:no_familiar and C3:bolster, all cases $p < 0.0001$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.841$, C1:familiar and C3:bolster, $p = 1.000$, and C2:no_familiar and C3:bolster, $p = 0.771$.

There was a significant difference between conditions at Session 2, $F(3,100) = 3.550, p = 0.017, r = 0.096$. As shown in Figure 4.24 on page 154, the SCR for C1:familiar was lower compared to C2:no_familiar, $p = 0.028$. There were no significant differences between C1:familiar and C3:bolster, $p = 0.710$, C1:familiar and C4:rest, $p = 1.000$, C2:no_familiar and C3:bolster, $p = 1.000$, C2:no_familiar and C4:rest, $p = 0.060$ and C3:bolster and C4:rest,

$p = 1.000$.

Heart rate

A mixed between-within subjects analysis of variance showed there was a significant effect of time on HR, $F(2.773, 266.25) = 18.47, p < 0.0001, r = 0.161$, and no significant effect of conditions, $F(3, 96) = 2.233, p = 0.089, r = 0.065$. There was significant interaction between condition and time on HR, $F(8.320, 266.25) = 3.566, p < 0.0001, r = 0.100$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. There was no significant change in HR from pre-measurement (t1) to Session 1 (t2), $p = 0.316$. The HR showed a significant increase from pre-measurement (t1) to post-measurement (t4), $p < 0.0001$. The HR showed a significant increase from Session 1 (t2) to Session 2 (t3), $p = 0.001$. (Refer to Figure 4.18). The HR showed no significant difference in HR from Session 2 (t3) to post-measurement (t4), $p = 1.000$.

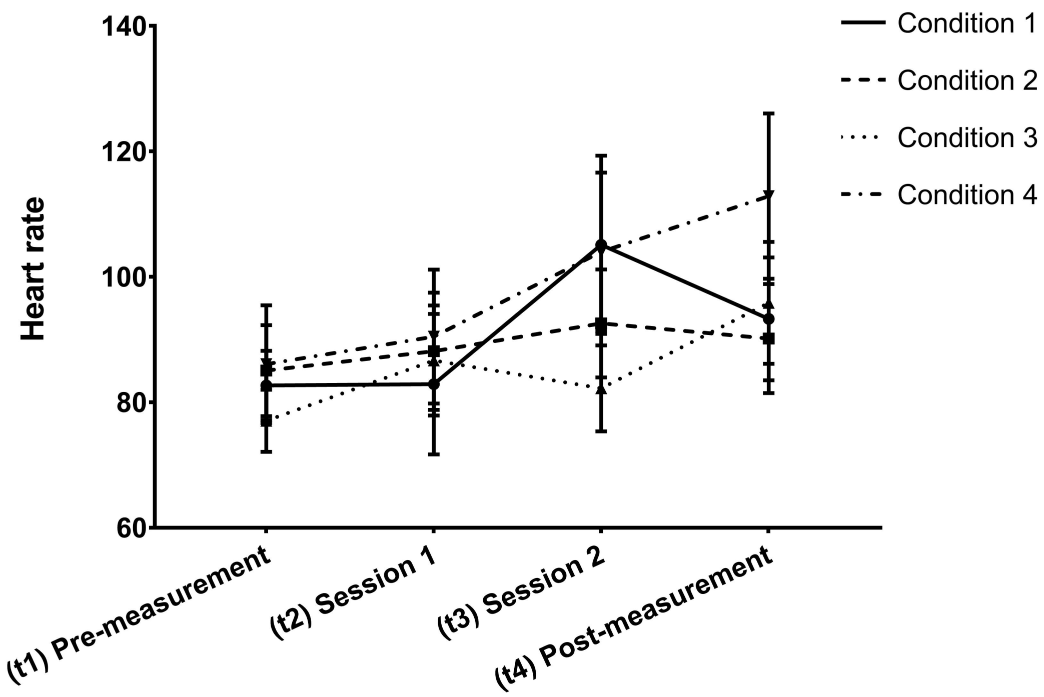


Figure 4.18: The mean with 95% confidence interval for heart rate to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in HR. The new variables are HR_t2mt1 = Session 1 minus pre-measurement, HR_t3mt2 = Session 2 minus Session 1 and HR_t4mt3 = post-measurement minus Session 2. Additional analysis, HR_t4mt1 = post-measurement minus pre-measurement.

There was no significant difference between conditions for HR_t2mt1, $F(3,97) = 0.855, p = 0.467, r = 0.026$.

There was a significant difference between conditions for HR_t3mt2, $F(3,98) = 5.288, p = 0.002, r = 0.139$. As can be seen in Figure 4.19, the increased HR was significantly higher in C1:familiar compared to C3:bolster, $p = 0.002$. The C3:bolster showed a small decrease in HR. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.051$, C2:no_familiar and C3:bolster, $p = 1.000$, C1:familiar and C4:rest, $p = 1.000$, C2:no_familiar and C4:rest, $p = 1.000$ and C3:bolster and C4:rest, $p = 0.097$.

There was a significant difference between conditions for HR_t4mt3, $F(3,98) = 4.717, p = 0.004, r = 0.146$. As is apparent in Figure 4.19, the HR was significantly lower in C1:familiar compared to C3:bolster, $p = 0.013$ and in C1:familiar compared to C4:rest, $p = 0.011$. There was a decrease trend for HR in C1:familiar and C2:no_familiar, and increase trend for HR in C3:bolster and C4:rest. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.896$, C2:no_familiar and C3:bolster, $p = 0.530$, C2:no_familiar and C4:rest, $p = 0.503$ and C3:bolster and C4:rest, $p = 1.000$.

There was a significant difference between conditions for HR_t4mt1, $F(3,97) = 5.024, p = 0.003, r = 0.134$. As shown in Figure 4.19, the increased HR was significantly lower in C1:familiar compared to C4:rest, $p = 0.040$ and in C2:no_familiar compared to C4:rest, $p = 0.020$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 1.000$, C1:familiar and C3:bolster, $p = 1.000$, C2:no_familiar and C3:bolster, $p = 0.051$, and C3:bolster and C4:rest, $p = 1.000$.

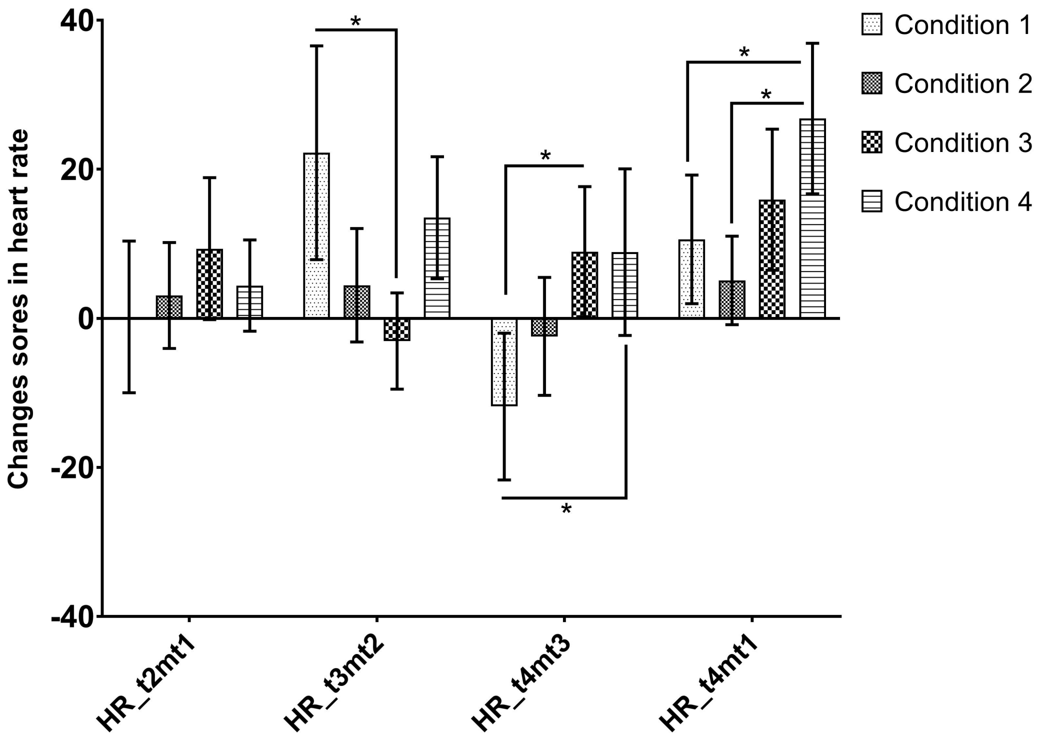


Figure 4.19: The mean with 95% confidence interval for changes in scores of heart rate to each condition. $*p < 0.05$

Session 1 (t2) and Session 2 (t3) There was no significant difference between conditions at Session 1, $F(3,100) = 0.425, p = 0.736, r = 0.013$. Although there was no significant difference, as can be seen in Figure 4.24 on page 154, the HR in C1:familiar was lower compared to other conditions.

There was a significant difference between conditions at Session 2, $F(3,98) = 3.896, p = 0.011, r = 0.107$. As is apparent in Figure 4.24 on page 154, the HR was higher in C1:familiar compared to C3:bolster, $p = 0.023$. The HR for C3:bolster was lower compared to C4:rest, $p = 0.035$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.601$, C2:no_familiar and C3:bolster, $p = 1.000$, C1:familiar and C4:rest, $p = 1.000$, and C2:no_familiar and C4:rest, $p = 0.804$.

Standard of all normal RR intervals

A mixed between-within subjects analysis of variance ANOVA showed there was a significant effect of time on SDNN, $F(2.652, 265.21) = 29.63, p < 0.0001, r = 0.229$, and no significant effect of conditions, $F(3, 100) = 2.018, p = 0.116, r = 0.057$. There was significant interaction between condition and time on SDNN, $F(7.956, 266.21) = 2.687, p = 0.008, r = 0.075$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. As evident in Figure 4.20, there was a significant increase in SDNN for pre-measurement (t1) to Session 1 (t2), $p < 0.0001$ and pre-measurement (t1) to post-measurement (t4), $p < 0.0001$. However, there was no significant change from Session 1 (t2) to Session 2 (t3), $p = 0.196$ and Session 2 (t3) to post-measurement (t4), $p = 1.000$.

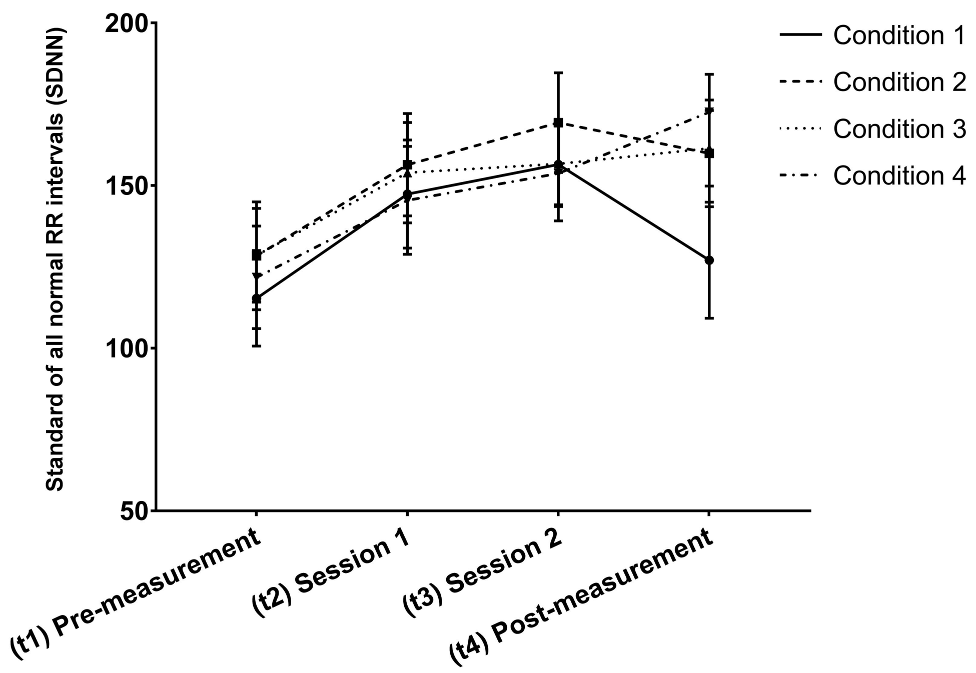


Figure 4.20: The mean with 95% confidence interval for SDNN to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in SDNN. The new variables are $SDNN\_t2mt1$ = Session

1 minus pre-measurement, $SDNN\_t3mt2$ = Session 2 minus Session 1 and $SDNN\_t4mt3$ = post-measurement minus Session 2. Additional analysis, $SDNN\_t4mt1$ = post-measurement minus pre-measurement.

There was no significant difference between conditions for $SDNN\_t2mt1$, $F(3,100) = 0.141, p = 0.935, r = 0.004$ and $SDNN\_t3mt2$, $F(3,100) = 0.306, p = 0.821, r = 0.009$. As is apparent in Figure 4.21, all conditions showed an increase in SDNN from pre-measurement (t1) to Session 1 (t2). All conditions showed no change in SDNN from Session 1 (t2) to Session 2 (t3).

There was a significant difference between conditions for $SDNN\_t4mt3$, $F(3,100) = 9.829, p < 0.0001, r = 0.228$ As evident in Figure 4.21, there was a decrease trend in C1:familiar and C2:no_familiar, and an increase trend in C3:bolster and C4:rest. The decreased SDNN showed a significant decrease in C1:familiar compared to C3:bolster, $p = 0.002$, and C1:familiar compared to C4:rest, $p < 0.0001$, and a significant decrease in C2:no_familiar compared to C4:rest, $p = 0.018$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.204$, C2:no_familiar and C3:bolster, $p = 0.775$, and C3:bolster and C4:rest, $p = 0.803$.

There was a significant difference between conditions for $SDNN\_t4mt1$, $F(3,96) = 5.387, p = 0.002, r = 0.144$. As shown in Figure 4.21, the increased SDNN was significantly lower in C1:familiar compared to C4:rest, $p = 0.022$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.804$, C1:familiar and C3:bolster, $p = 0.654$, C2:no_familiar and C3:bolster, $p = 1.000$, C2:no_familiar and C4:rest, $p = 0.866$ and C3:bolster and C4:rest, $p = 1.000$. As can be seen in Figure 4.21, the C4:rest showed a larger increase compared to other conditions, although not significant.

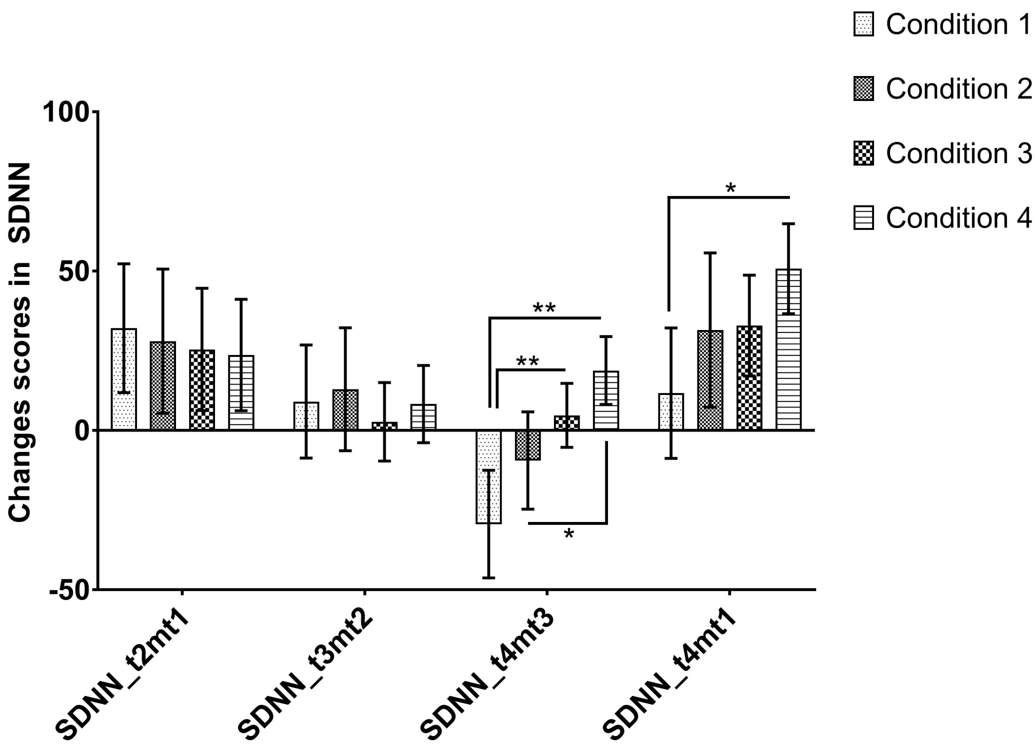


Figure 4.21: The mean with 95% confidence interval for changes in scores of SDNN to each condition. $*p < 0.05, **p < 0.001$

Session 1 (t2) and Session 2 (t3) There was no significant difference between conditions at Session 1, $F(3,100) = 0.446, p = 0.720, r = 0.013$.

There was no significant difference between conditions at Session 2, $F(3,100) = 1.074, p = 0.364, r = 0.031$. Although there was no significant difference, as can be seen in Figure 4.25 on page 155, for both sessions, the SDNN in C2:no_familiar was higher compared to other conditions.

High frequency

A mixed between-within subjects analysis of variance ANOVA showed there was a significant effect of time on HF, $F(3,300) = 10.51, p < 0.0001, r = 0.095$, and no significant effect of conditions, $F(3,100) = 1.117, p = 0.346, r = 0.032$. There was no significant interaction between condition and time on HF, $F(9,300) = 0.993, p = 0.446, r = 0.029$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. There was a significant increase in HF for pre-measurement (t1) to Session 1 (t2), ($p < 0.0001$ and pre-measurement (t1) to post-measurement (t4), $p < 0.0001$. (Refer to Figure 4.22). However, there was no significant change from Session 1 (t2) to Session 2 (t3), $p = 0.196$ and Session 2 (t3) to post-measurement (t4), $p = 1.000$.

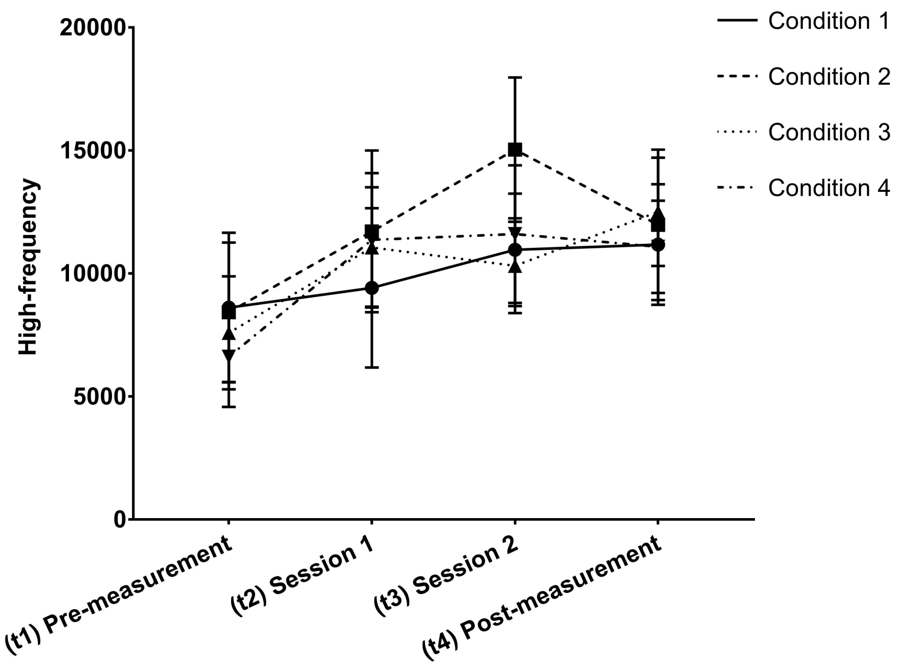


Figure 4.22: The mean with 95% confidence interval for high-frequency to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in HF. The new variables are HF_t2mt1 = Session 1 minus pre-measurement, HF_t3mt2 = Session 2 minus Session 1 and HF_t4mt3 = post-measurement minus Session 2, and HF_t4mt1 = post-measurement minus pre-measurement.

There was no significant difference between conditions for HF_t2mt1, $F(3,100) = 0.757, p = 0.521, r = 0.022$, HF_t3mt2, $F(3,100) = 1.190, p = 0.317, r = 0.034$ and HF_t4mt3, $F(3,100) = 2.390, p = 0.073, r = 0.067$.

There was also no significant difference between conditions for HF_t4mt1, $F(3,100) = 0.383, p = 0.766, r = 0.011$, and HF_t3mt1, $F(3,100) = 1.325, p = 0.271, r = 0.038$. As is apparent in Figure 4.23, all conditions showed an increase in HF scores for HF_t2mt1 and HF_t4mt1. All conditions showed no change in HF scores for HF_t3mt2 and HF_t4mt3.

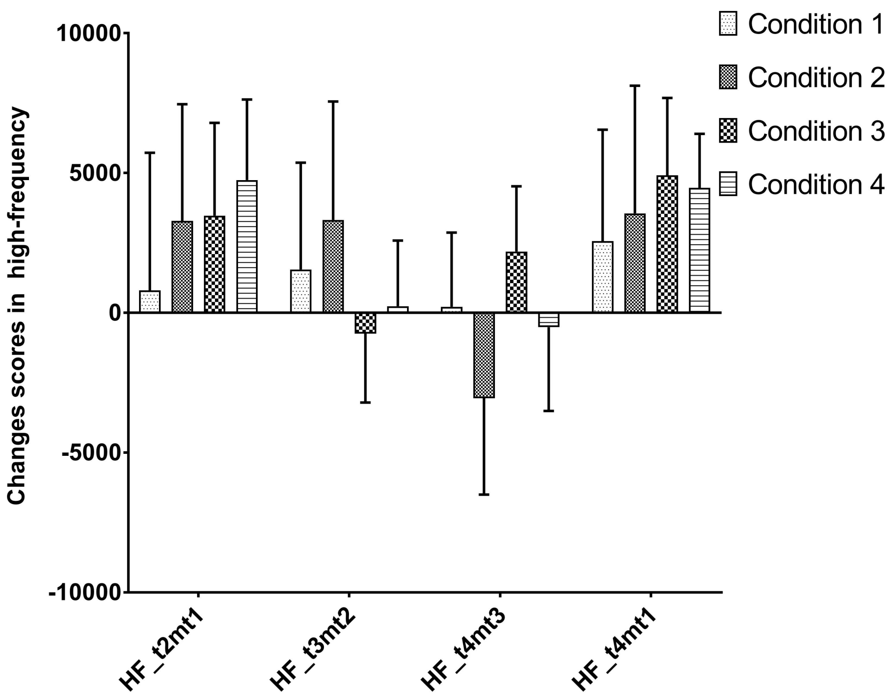


Figure 4.23: The mean with 95% confidence interval for changes in scores of high-frequency to each condition.

Session 1 (t2) and Session 2 (t3) A one-way between-subjects ANOVA showed there was no significant difference between conditions at Session 1, $F(3,100) = 0.507, p = 0.679, r = 0.015$. Although there was no significant difference, as can be seen in Figure 4.25 on page 155 , the HF in C1:familiar was lower compared to other conditions.

A one-way between-subjects ANOVA showed there was a significant difference between conditions at Session 2, $F(3,100) = 2.961, p = 0.036, r = 0.082$. As is apparent in Figure 4.25 on page 155 , the HF for C2:no_familiar was higher compared to C3:bolster, $p = 0.045$. There were no significant differences between C1:familiar and C2:no_familiar, $p = 0.123$, C1:familiar and C3:bolster, $p = 1.000$, C1:familiar and C4:rest, $p = 1.000$, C2:no_familiar and C4:rest,

$p = 0.299$ and C3:bolster and C4:rest, $p = 1.000$.

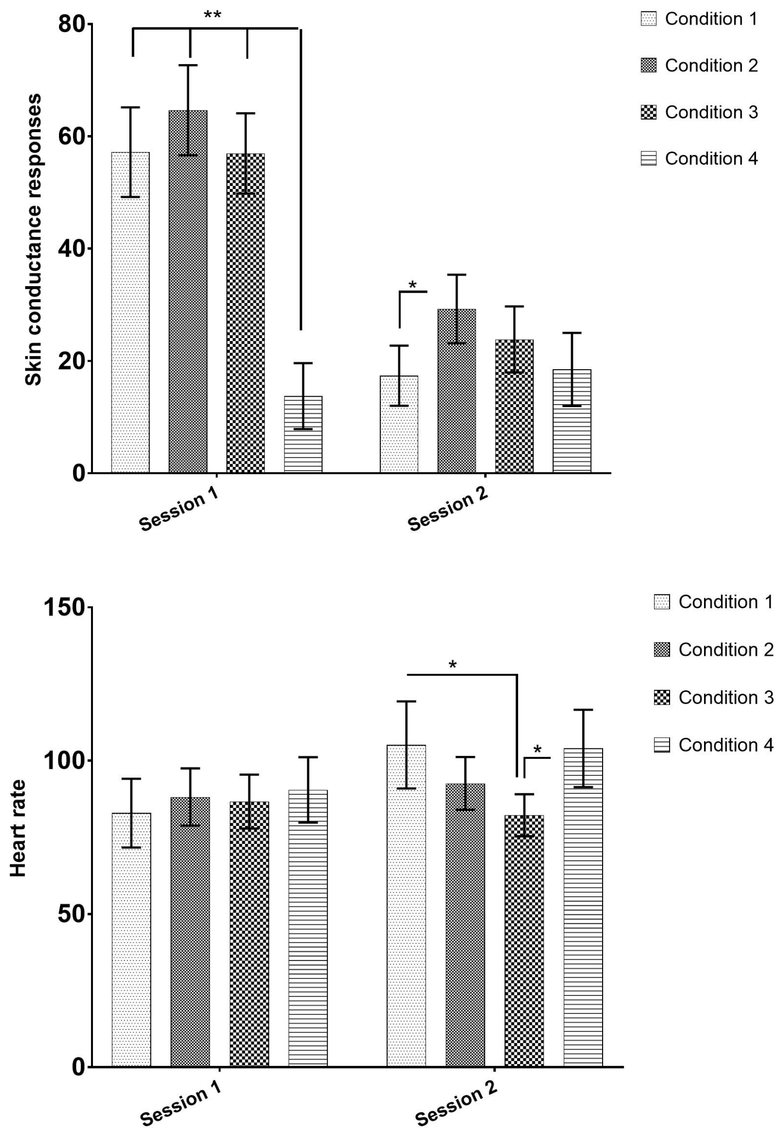


Figure 4.24: The mean with 95% confidence interval for Session 1 and Session 2 of skin conductance responses and heart rate to each condition. $*p < 0.05, **p < 0.001$.

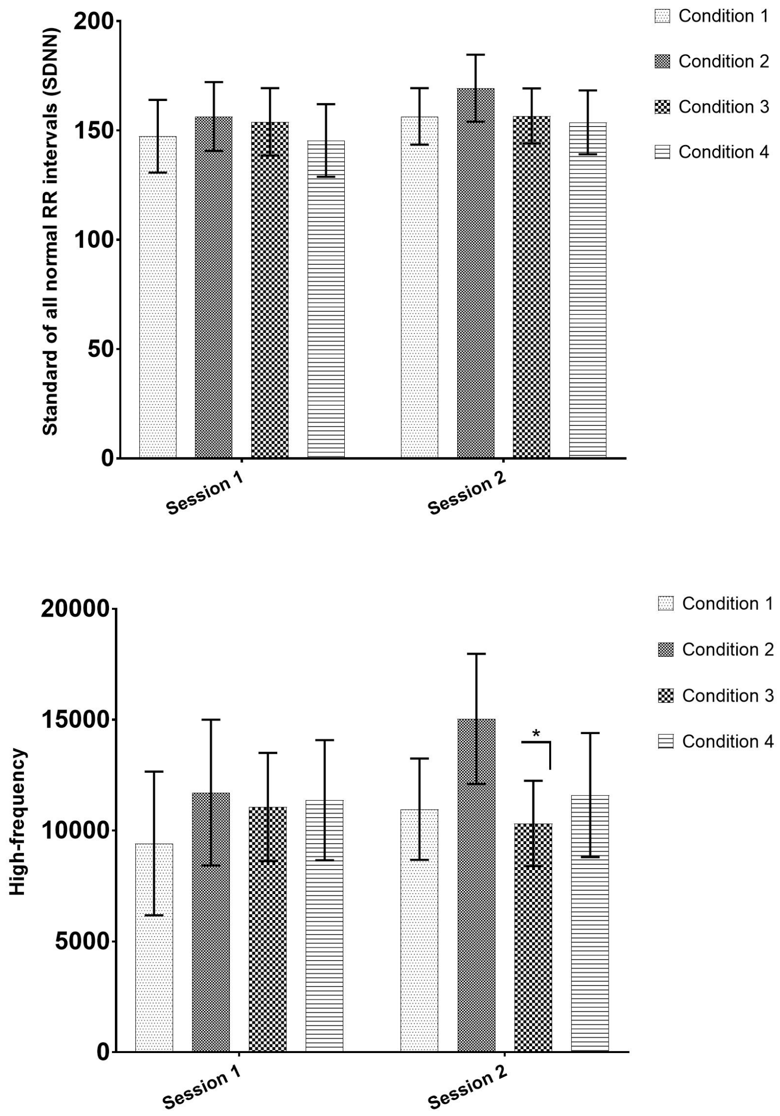


Figure 4.25: The mean with 95% confidence interval for Session 1 and Session 2 of SDNN and HF to each condition. $*p < 0.05$.

4.11 Summary of findings based on predictions

4.11.1 H1: Effects of conditions in Session 1

This section focuses on the findings for Hypotheses H1(a) and H1(b).

H1(a): Hypothesis for self-reported responses:

- There was no significant difference in self-reported stress across conditions, all conditions showed a low stress. (Refer to Figure 4.11, on page 136).
- There was a significant effect of conditions in self-reported negative affect. The negative affect was significantly lower in C2:no_familiar compared to C4:rest, and in C3:bolster compared to C4:rest. But, there was no significant difference between C1:familiar and C3:bolster, and between C1:familiar and C4:rest. (Refer to Figure 4.11, on page 136).
- There was a significant difference in self-reported positive affect between Paro conditions (C1:familiar and C2:no_familiar) compared to in C3:bolster and C4:rest. (Refer to Figure 4.11, on page 136).
- There was a significant difference between conditions in self-reported emotions. Self-reported positive emotion was rated significantly higher in Paro conditions (C1:familiar and C2:no_familiar) compared to bolster condition (C3:bolster) and rest condition (C4:rest). The positive emotions rated were happy, excited and relaxed. However, only excited was rated significantly higher in C1:familiar and C2:no_familiar, compared to C3:bolster. There were no differences in rated happy and relaxed across conditions. (Refer to Figure 4.13, on page 138, and Figure 4.12, on page 137.)
- Self-reported negative emotions were rated significantly lower in Paro condition (C1:familiar) compared to bolster condition (C3:bolster). There was no significant difference between C1:familiar and C2:no_familiar, and between C2:no_familiar and C3:bolster. The negative emotions rated were disgusted, bored and confused. Confused was rated significantly higher in C2:no_familiar and C3:bolster, compared to C1:familiar. (Refer to Figure 4.13, on page 138, and Figure 4.12, on page 137.)
- Therefore, *H1(a) is partially supported*, through differences in self-reported negative and positive affect, and self-reported negative and positive emotions.

Summary: It seems that interacting with an active Paro (in both familiar and unfamiliar session) results in a higher level of positive affect and positive emotions, and a lower level of negative affect, negative emotions, and self-reported stress.

H1(b): Hypothesis for physiological responses:

- There was no significant difference in levels of HR and HRV parameters for all conditions. The SCR were significantly higher in all three conditions except C4:rest. (Refer to Figure 4.25, on page 155, and Figure 4.24, on page 154.)
- Therefore, *H1(b) is not supported.*

Summary: It seems that all conditions create similar physiological responses in HR and HRV, but not in SCR, where the arousal is shown only in interactive conditions.

H1(c): Hypothesis for observed facial expressions:

- *Positive facial expression:* There were significant differences in observed positive between (C1:familiar and C2:no_familiar) and C3:bolster. It is clear that participants were showing positive emotions when given an option to interact freely with an active Paro, as indicated by greater number of positive expressions observed in C1:familiar and C2:no_familiar compared to the bolster condition (C3:bolster). Participants in C1:familiar and C2:no_familiar were observed to have a high frequency of smiling expressions compared to C3:bolster. (Refer to Figure 4.15, on page 141, and Figure 4.14, on page 140.)
- *Negative facial expression:* The hypothesis is supported through the difference in observed negative facial expression between (C1:familiar and C2:no_familiar) and C3:bolster. C3:bolster showed significantly higher frequency of negative expression compared to C1:familiar and C2:no_familiar. Participants in C3:bolster were observed to have a high frequency of bored expressions. (Refer to Figure 4.15, on page 141, and Figure 4.14, on page 140.)

- Therefore, *H1(c) is supported.*

Summary: It seems that a session with Paro create more positive emotions than a session stroking a bolster.

4.11.2 H2: Effects of conditions on stress responses

H2(a): Hypothesis for self-reported responses:

- Unexpectedly, there was no significant difference between the condition in self-reported stress, positive affect, and negative affect. The self-reported positive affect showed a decrease in all conditions. The results showed an increase in self-reported stress and negative affect for all conditions. (Refer to Figure 4.6, on page 129, Figure 4.8, on page 132, and Figure 4.10, on page 135.)
- Therefore, *H2(a) is not supported.*

H2(b): Hypothesis for physiological responses:

- *SCR and HR:* Study 2 did not find a difference in SCR between conditions, although the SCR did return to baseline. The HR, however, showed a further increase across conditions. Despite the increased HR, the HR in C1:familiar and C2:no_familiar were significantly lower than C4:rest. (Refer to Figure 4.17, on page 144, and Figure 4.19, on page 148.)
- *HF and SDNN:* Study 2 found a further increase across conditions for HF and SDNN, from baseline to final measurement. But, there was a significant difference between C1:familiar and C4:rest: the SDNN in C1:familiar were lower than C4:rest. (Refer to Figure 4.23, on page 153, and Figure 4.21, on page 151.)
- Therefore, *H2(b) is partially supported*, through differences between conditions in HR and SDNN.

- Summary: There is some indication here that a session with Paro can results in lower heart rate and increased SDNN. This suggests Paro can contribute to alleviate a stress response.

4.11.3 H3: Effects of familiarisation with Paro compared to no familiarity to Paro in Session 1

H3(a): Hypothesis for self-reported responses:

- There was no difference between C1:familiar and C2:no_familiar on self-reported stress, negative and positive affect, and positive and negative emotions. Therefore, the *H3(a) is not supported.*
- But, based on self-reported emotions, 'confused' was rated significantly lower when interacting with active Paro in C1:familiar compared to C2:no_familiar, when participants had a time to get familiar with the Paro before the experiment. (Refer to Figure 4.12, on page 137.)

H3(b): Hypothesis for physiological responses:

- There was no difference between C1:familiar and C2:no_familiar on skin conductance responses, heart rate and HRV parameters, *H3(b) is not supported.* (Refer to Figure 4.17, on page 144, Figure 4.19, on page 148, Figure 4.23, on page 153, and Figure 4.21, on page 151.)

H3(c): Hypothesis for observed facial expressions:

- There was no difference between C1:familiar and C2:no_familiar on observed positive and lower negative facial expressions, *H3 (c) is not supported.* (Refer to Figure 4.15, on page 141, and Figure 4.14, on page 140.)

4.11.4 H4: Effects of familiarisation with Paro compared to no familiarity to Paro on stress responses

H4(a): Hypothesis for self-reported responses:

- There was no difference between C1:familiar and C2:no_familiar on self-reported stress and negative affect and positive affect, *H4(a) is not supported.* (Refer to Figure 4.6, on page 129, Figure 4.8, on page 132, and Figure 4.10, on page 135.)

H4(b): Hypothesis for physiological responses:

- There was no difference between C1:familiar and C2:no_familiar on SCR, HR and HRV parameters, *H4(b) is not supported.* (Refer to Figure 4.17, on page 144, Figure 4.19, on page 148, Figure 4.23, on page 153, and Figure 4.21, on page 151.)

Summary: Apart from the self-reported emotion of confusion, there is no indication here that 2 minutes familiarisation with the Paro had an the effect on physiological and self-reported measures.

4.12 Discussion

The results will be discussed in two parts, in order to focus on *(Part 1:)* the results obtained during Session 1, and *(Part 2:)* the results obtained in the form of changes in scores between baseline and post-measurement, in order to look at the evidence of the stress-reducing effects of Session 1 on subsequent induced stress.

4.12.1 Part 1: Effects of conditions during Session 1

Considering the results obtained in Session 1 (at t2), in sum, only one of the three hypotheses in H1 was definitely supported. *H1(a) was partially supported*, with a small effect of conditions on self-reported positive and negative affect, and self-reported emotions. *H1(b) was not supported*, as no difference between physiological responses were found between any of the conditions. *H1(c)* was supported, since there was a significant difference in observed positive and negative facial expressions between (C1:familiar and C2:no_familiar) and C3:bolster. *H3(a, b, c)* was not supported, as there was no significant difference between C1:familiar and C2:no_familiar in Session 1 (at t2) for any of the measures.

A comparison between the findings of Study 1 and Study 2

This section focuses on a comparison between the findings of Study 1 and Study 2 during the intervention session.

Study 1 investigated the effects of interacting with Paro following a stress-inducing session. By contrast, Study 2 investigated the effects of interacting with Paro before a stress-inducing session. The change in ordering of the intervention session and noise session in Study 2 was carried out, so that the effects of interacting with Paro without the influence of previous stress inducing noise could be better understood. It should be noted that any effects that occurred during the intervention session in Study 1 may have been influenced by the previous noise session.

It should be noted that in Study 2, new measures were added to give more information about what happened during the intervention session and its effects on stress responses. The new measures were time length of intervention session, self-reported emotions, and physiological measures, heart rate (HR) and heart rate variability (HRV) parameters. But, the results from the new measures will not be discussed in this sub-section. Despite these differences, when comparing the results of Study 1 and Study 2, there are some consistent findings in the self-reported and skin conductance measures taken during the intervention session.

Firstly, we will consider the results of skin conductance response (SCR) in Study 1 and Study 2. In Study 2, it was found that the SCR were higher in all three conditions (involving interaction with a Paro or stroking a bolster) compared to the resting alone condition. *These results are similar to the results of SCR in Study 1, where the SCR were also higher in all three conditions that involved interaction with Paro as compared to the rest condition.*

In Study 1, it was suggested that the arousal of SCR was associated with positive moods, as shown in the highly rated positive affect in all the three Paro conditions. However, high positive affect was also found in C4:inactive_rest, where participants were asked to rest with the presence of an inactive Paro. So, based on the high arousal in three conditions, and the highly rated positive moods in all conditions in Study 1, it was further suggested that this could be due to either (i) the cute appearance of Paro influencing the increased positive mood (either when the Paro was active or inactive), or (ii) the effect of stroking a nice object covered with fur may have influenced the moods and arousal of SCR, or both. Therefore, in Study 2, the cute appearance of Paro was eliminated in C3:bolster (replaced by a bolster) and in C4:rest (participants were left alone). This makes it possible to examine the effects of stroking on physiological arousal and moods without the presence of a Paro.

The results of Study 2 could be interpreted as showing that the increased SCR (from baseline to Session 1) was related to positive arousal when participants were asked to interact with Paro (either in C1:familiar and C2:no_familiar). However, the increased SCR in C3:bolster (when stroking a bolster) is less likely to be related to positive emotions because the positive affect was significantly lower in C3:bolster compared to C1:familiar and C2:no_familiar.

This would mean that, when comparing between Study 1 (C3:inactive_stroke) and Study 2 (C3:bolster), the high SCR and high positive affect in Study 1 (in C3:inactive_stroke) are not only related to stroking an inactive Paro, but are also related to the presence of Paro. This is because in Study 2, when the Paro was not present, but was replaced with bolster, the SCR was high but the positive affect was low in C3:bolster (stroking the bolster). This suggests that in C3:bolster, the higher arousal SCR was due to effects of stroking the bolster.

Table 4.2 summarises results obtained at Session 1 (1) on self-reported responses and SCR, based on data observed in graphs and the statistical effects of conditions.

Table 4.2: Effects of conditions in Session 1

At Session 1	C1	C2	C3	C4
Self-reported stress	Low (-)	Low (-)	Low (-)	Low (-)
Self-reported positive affect	High (*C3 C4)	High (*C3 C4)	Low (*C1 C2)	Low (*C1 C2)
Self-reported negative affect	Low (-)	Low (*C4)	Low (*C4)	High (*C2 C3)
SCR	High (*C4)	High (*C4)	High (*C4)	Low (*C1 C2 C3)

Note. Based on results of self-reported responses obtained at Session 1 (t2). High indicates high scores observed in condition; Low indicates low scores observed in condition; an asterisk indicates the condition was significantly higher or lower compared to a condition or conditions; a hyphen (-) indicates no significant difference between any of conditions.

However, in Study 2, based on the results on self-reported negative affect and stress during intervention session (from baseline to Session 1), the results showed that the bolster does have the potential to reduce self-reported negative affect and stress. This is because no significant difference was found between Paro conditions (C1:familiar and C2:no_familiar) and the bolster condition (C3:bolster). The results for these conditions all showed low self-reported negative affect and stress at Session 1 (at at2).

Furthermore, in Study 2, when comparing the effect of Paro conditions and the resting alone condition on self-reported stress during intervention session, it was found that the Paro condition (C2:no_familiar) had a significantly lower self-reported stress (from baseline (t1) to Session 1 (t2)) than C4:rest (where no Paro was present). But, in Study 1, a significantly lower self-reported stress (from noise session (t2) to intervention session (t3)) was found in all conditions. *This also suggests that the presence of an inactive Paro in the room (Study 1*

(C4:inactive_rest)) may also influence the levels of stress compared to resting alone without the presence of Paro (Study 2 (C4:rest)).

Table 4.3 summarises results obtained in changes in scores between baseline and Session 1 (t2mt1) on self-reported and physiological responses.

Table 4.3: Effects of conditions on changes in scores between baseline and Session 1 (t2mt1)

t2mt1	C1	C2	C3	C4
Levels of stress	↓	↓ (*C4)	↓	↓
PA	↑(*C4)	↑ (*C3 C4)	↓ (*C2)	↓ (*C1 C2)
NA	↓ (*C4)	↓ (*C4)	↓	↓
SCR	↑ (*C4)	↑ (*C4)	↑ (*C4)	↓ (*C1 C2 C3)
HR	-	-	-	-
SDNN	↑	↑	↑	↑
HF	↑	↑	↑	↑

Notes. ↑ indicates a significant increase from baseline to Session 1; ↓ indicates a significant decrease from baseline to Session 1; a hyphen (-) indicates no change in scores from baseline to Session 1. ↑ and an asterisk indicates the condition was significantly higher compared to a condition or conditions; ↓ and an asterisk indicates the condition was significantly lower compared to a condition or conditions.

Effects of familiarisation with Paro at Session 1

Another suggestion in Study 1 was that SCR arousal was related to the stress created as a result of interaction with unfamiliar object. Therefore, in Study 2, a familiarisation session before the experiment was added (C1:familiar), in order to compare its effects to Paro with no familiarity.

No difference was found in self-reported and physiological responses when comparing between C1:familiar and C2:no_familiar during Session 1 (at t2), and *H3 was not supported.* But, both of the conditions (C1:familiar and C2:no_familiar) showed a high self-reported positive and low self-reported negative affect and stress. At the same time, both conditions showed high arousal SCR, and similar levels HR and HRV parameters, which are suggested to be related to excitement and relaxation.

This result, in contrast to a study by Okita (2013), the author found no reduction in anxiety in hospitalized children after they had a short session with an active Paro (compared to baseline). The author suggested that the lack of changes in levels of anxiety was related to unfamiliarity with Paro. However, in the Okita (2013) study, different samples of participants and settings were used. Okita (2013) recruited children that had been warded for a certain illness, although illness was not reported by the author. *Children may interact differently to Paro compared to young healthy adults.* For example, based on a review by Leite et al. (2013), the authors mentioned that healthy young adults tended to accept playing with the Paro compared to elderly people with dementia (Turkle et al. 2006). Turkle et al. (2006) mentioned that when the elderly are introduced to pet robot, they showed anxiety and concern about how the pet robot will behave. Additionally, Robinson et al. (2013b) also stated that difficulties in physical movement might have been one of factor for the lack of interaction with the pet robot.

Furthermore, a natural setting such as hospital could be more stressful than being exposed to noise in laboratory-setting. But, the effects of familiarisation on stress reduction will be discussed further in Part 2, section 4.12.2, page 176.

Even though hypothesis (H3) was not supported in Study 2, there was an effect of familiarity on self-reported emotions (at t2). Participants reported that they felt more confused when interacting with Paro in C2:no_familiar compared to C1:familiar. The low rate of reported 'confused' in C1:familiar is presumably because participants understood how to interact with Paro (after the 2 minutes familiarisation session with Paro) and already knew how Paro would respond to their interaction. Despite a lack of understanding about Paro in C2:no_familiar, it was showed that the emotions of happy and excited were rated the highest in C2:no_familiar compared to C1:familiar. In addition, both of the Paro conditions showed significantly higher levels of reports of 'excitement' compared to C3:bolster. This finding is also *supported by H1(c)* because no difference was found between C1:familiar and C2:no_familiar in observed positive and negative facial expression. It seems that familiarity reduced the novelty effect of Paro, but did not reduce the effects of Paro on positive moods. *This finding is in line with previous long-term studies of Paro that mentioned that participants maintained their positive* *interaction with Paro even after a few sessions with the robot* (Sabanovic et al. 2013, Birks et al. 2016).

A comparison between baseline (t1) and Session 1 (t2) on heart rate and heart rate variability parameters

In this sub-section, some details about what happened during Session 1 when compared to the baseline (t2mt1) on heart rate (HR) and heart rate variability (HRV) parameters are discussed.

HRV:

In addition to findings of SCR in Study 2, all the HRV parameters i.e. SDNN and HF were significantly increased in all conditions. The increased HRV in all conditions may indicate a greater level of relaxation (when stroking a Paro or a bolster, or resting) or excitement (when playing with Paro). This is related to the domination of the parasympathetic activity (Quintana et al. 2012, Castaldo et al. 2015), as also reflected in self-reported stress, negative and positive affect.

HR:

Interestingly, no change in HR was found in any of the conditions from baseline to Session 1 (t2mt1). This may be related to parasympathetic activity, that helps to maintain the resting heart rate (as compared changes from baseline and Session 1) (Park et al. 2018a). The findings of Study 2 are partially different to previous studies of Paro that investigated the effects of Paro on HR and HRV (Mitsui et al. 2001*a*,*b*). In a previous study by Mitsui et al. (2001*a*), participants were assigned to four conditions where they were asked to stroke (1) a column covered with Paro's fur, or (2) a column covered in vinyl, (3) inactive Paro and, (4) active Paro. They found a significantly greater increase in HR from baseline to a session when participants stroked a column covered with Paro's fur compared to stroking a column covered with vinyl material, and compared to playing with active Paro. They found no change in HR from baseline when participants interacted with active Paro. It was suggested that the low HR was associated with concentration when playing with active Paro.

The results of Study 2 in HR are partially consistent with the study by Mitsui et al. (2001*a*), since the HR did not show any changes from baseline to the interacting with active Paro session, and stroking a bolster. This may suggest that the Paro and bolster conditions may have had similar effects on HR in Study 2.

It is possible to find some reasons for the differences in findings between Study 2 and Mitsui et al. (2001*a*) on HR. They may be related to the within-subject experiment design and unintentional personal bias when stroking a similar object to Paro in Mitsui et al. (2001*a*) study. It possibly affected the physiological responses, where participants anticipated about how Paro would interact or preferred to stroke Paro because they have been in a session with Paro before, hence the increased HR (in the column covered with Paro's fur condition). It should be noted that the authors (Mitsui et al. (2001*a*)) used a random sequence of the four conditions for each participant, to avoid bias. However, as mentioned in a review by (Pannucci & Wilkins 2010), personal bias during research cannot be fully avoided.

Additionally, the different effects on physiological responses may also be related to the different softness of Paro's fur and the bolster used in Study 2. However, it is beyond the scope of Study 2, to look at the effect of different softness of Paro and bolster.

HRV:

Furthermore, Mitsui et al. (2001*a*) also found that the HRV showed a greater increase when participants played with an active Paro compared to other conditions. Mitsui et al. (2001*a*) predicted that the HRV would be lowered when interacting with Paro, as they used a ratio of parasympathetic and sympathetic activity as a HRV measure (i.e. LF/HF). Mitsui et al. (2001*a*) also suggested that despite the increase in HRV, there were improvements in self-reported moods in the active Paro condition. Based on a review on HRV by Castaldo et al. (2015), the HRV ratio is related to domination of sympathetic activity. However, there were some studies that have shown an increase in HRV ratio during relaxation therapy (Sripongngam et al. 2015, Seifert et al. 2018), and during mild physical exercise (Sakakibara et al. 1994). Seifert et al. (2018) suggested that the increase of HRV ratio is due to increases in baroreflex activity which is related to relaxation. The baroreflex reacts to any changes in blood pressure, and it helps to maintain the pressure to a resting level.

Even though Study 2 used different parameters of HRV, the results are partially consistent with Mitsui et al. (2001*a*). The increased HRV parameters in Paro conditions in Study 2, were not only related to relaxation, but also associated with excitement.

Conclusion of Part 1

Overall, when considering the results obtained in Session 1, it was noted that interacting with an active Paro induced excitement. The effects reflected in increased self-reported positive affect and positive emotions, and increased SCR and HRV parameters, from baseline to Session 1. At the same time, there is also an indication of a greater reduction in self-reported stress and negative affect, and no changes in HR, from baseline to Session 1, when interacting with Paro and stroking a bolster were compared to the resting alone condition (C4:rest).

In Part 2, there follows a discussion about whether these effects can help to protect participants from subsequent stress exposure, whether the high positive and negative affects have a lasting effect during or after the stress session, and whether the effects could help in faster recovery from the induced stress after the stress session ended.

4.12.2 Part 2: Effects of conditions on subsequent induced stress

This section will continue to discuss the findings of Study 2 at the end of experiment when compared to baseline (at3mat1 or t4mt1), based on hypotheses: H2 and H4.

In order to induce mild stress, participants were exposed to a noise. In Study 1, it was found that the noise reliably induced mild stress. The stress response was indicated by significantly increased self-reported negative affect and stress, and skin conductance response (SCR), from baseline to the stress exposure session. In Study 2, the noise was slightly modified, the length of noise exposure was increased from 3 to 5 minutes. An additional sound of babies crying was added to the previous sound used in Study 1, in order to increase the levels of stress responses. Based on the results of Study 1, and past research on effects of noise on stress responses (Basner et al. 2014, Münzel et al. 2014a), the use of this type of noises should induce stress.

Overall findings

H2(a) was rejected, as there was no difference between conditions in all self-reported responses. However, *H2(b) was partially supported* as there was a significant effect of conditions in heart rate (HR) and standard of all normal RR intervals (SDNN). *H4(a,b) was not supported*, as there was no difference between C1:familiar and C2:no_familiar in all self-reported and physiological responses.

Unexpectedly, the self-reported stress and negative affect increased, and positive affect decreased, and HR and heart rate variability (HRV) parameters also increased, from baseline to final measurement (at3mat1 or t4mt1). Only SCR returned to baseline at the end of experiment. Overall, the results indicate that none of the conditions has the effect of reducing the self-reported responses and HR and HRV parameters as responses to noise exposure.

Table 4.4 and Table 4.5 summarises results of changes in scores in physiological and self-reported responses, respectively.

Table 4.4: Effects of conditions in changes scores on physiological responses

Changes in scores	SCR	HR	SDNN	HF
t2mt1	↑ (*)	ns (-)	↑ (-)	↑ (-)
t3mt2	↓ (*)	↑ (*)	ns (-)	ns (-)
t4mt3	ns (*)	ns (*)	ns (*)	ns (-)
t4mt1	ns (-)	↑ (*)	↑ (*)	↑ (-)

Table 4.4 continued from previous page

Changes in scores	SCR	HR	SDNN	HF

Notes. ↑ indicates a significant increase in changes scores; ↓ indicates a significant decrease in changes scores; ns indicates no change in changes scores; a hyphen (-) indicates there was no effect of conditions in changes scores; An asterisk indicates there was an effect of conditions in changes scores;

Table 4.5: Effects of conditions in changes scores on psychological responses

Changes in scores	Stress	PA	NA
at2mat1	↓ (*)	↑ (*)	↓ (*)
at3mat2	↑ (-)	↓ (*)	↑ (-)
at3mat1	↑ (-)	↓ (-)	↑ (-)

Notes. ↑ indicates a significant increase in changes scores; ↓ indicates a significant decrease in changes scores; ns indicates no change in changes scores; a hyphen (-) indicates there was no effect of conditions in changes scores; An asterisk indicates there was an effect of conditions in changes scores;

Possible reasons for lack of stress reduction at the end of experiment, compared to baseline

There are no previous studies of Paro that looked at the skin conductance response (SCR) and heart rate variability (HRV) responses for the effects of interaction with Paro on subsequent induced stress in laboratory-settings study. Therefore, the relationship between the present findings and animal therapy will be discussed. The results partially corroborate a previous study of the effect of animal therapy on subsequent induced stress by Fiocco & Hunse (2017).

In the Fiocco & Hunse (2017) study, participants were randomly allocated to interact with a dog for a few minutes, or to sit alone in the room (Session 1). After that, the dog was removed from the room, and they were asked to complete a stressful computerized task (Session 2). The authors found an increased trend of self-reported negative affect, and a decreased trend in self-reported positive affect in the dog condition compared to the resting alone condition, at the end of the experiment (compared to baseline). At the same time, they also found an increased SCR in both of the conditions, but a lower increase in SCR in the dog condition compared to the alone condition. This suggests that in the Fiocco & Hunse (2017) study, despite the high negative affect in the dog condition, the physiological stress responses (SCR) decreased.

According to Fiocco & Hunse (2017), the contradictory results between self-reported and physiological responses might have been influenced by an external variable such as personal bias, which may be related to pet ownership.

However, in the case of Study 2, a possible explanation for the lack of stress reducing effects of Paro in Study 2, in comparison to the Fiocco & Hunse (2017) study is that the authors used a different stress induction method. *In the case of Study 2, in terms of stressor, it could be that the noise exposure was too stressful compared to the stressor used* in Fiocco & Hunse (2017) study. Hence, participants might not have been able or might have taken a longer time to cope with the induced stress. As a consequence, there was increased self-reported stress and negative mood, and decreased positive affect further than a baseline at the end of experiment.

Another possible reason for no stress reduction in Study 2, could be linked to participants' pre-existing stress levels. It could be assumed that some of the participants have high levels of pre-existing stress, which are related to noise exposure. Hence, reducing the effects of the previous intervention session on noise exposure. There is a study by Jordan et al. (2016), who suggested that participants perceived their high levels of stress related to a long-term exposure to noise. However, Study 2 did not collect information regarding the participants' long term stress. But, it was acknowledged that participants had low negative moods at the beginning of the study in all conditions.

A comparison between Study 1 and Study 2 on findings of SCR, at the end of experiment compared to baseline

As mentioned earlier, interestingly, only the skin conductance response (SCR) returned to baseline at the end of the experiment across the four conditions.

The results of Study 2 for SCR were similar to the results of SCR in Study 1, where at the end of the experiment, the SCR returned to baseline for all conditions (See Figure 3.13, page 94). But, in Study 1, one of suggestions was that the presence of Paro in all conditions may have influenced the reduction in SCR. However, in Study 2, it was expected that Paro conditions would have a stronger effect to reduce the stress responses compared to other conditions where Paro was not present. However, unexpectedly, no effect of conditions was found at the end of experiment (when compared to baseline).

But, it is worth noting that, there was a trend showing a small decrease in SCR in C1:familiar and C2:no_familiar compared to C3:bolster and C4:rest, which showed a small increase in SCR at the end of experiment when compared to baseline (See Figure 4.17, page 144). This may suggest that interacting with Paro had some effect on reducing the stress responses, in terms of SCR. This finding is consistent with the study by Petersen et al. (2017), where the authors also found a reduction in stress in terms of SCR. But, in the study by Petersen et al. (2017), participants had a 20 minutes session with Paro for a few weeks. This suggests that even a short session with Paro in Study 2 had a similar benefit to long-term intervention, in terms of reducing stress responses.

Effects of conditions in accelerating faster recovery in physiological stress responses

In this section, an explanation about a faster recovery from stress reactivity will be discussed. As mentioned earlier, there are no past studies of Paro that used heart rate (HR) and heart rate variability (HRV) parameters to measure responses of the effects of interacting with Paro prior to stress induction in laboratory-settings study.

Nevertheless, based on the findings on effects of stress induction on physiological responses, the findings of Study 2 are partially consistent with previous studies of the effects of acute noise exposure on physiological responses. The previous studies showed that the acute noise increased HR (Mahmood et al. 2006, Sim et al. 2015), and increased SDNN (Sim et al. 2015, Walker et al. 2016), after the noise exposure ended as compared to baseline. According to (Walker et al. 2016), the increased HRV parameters (i.e. standard of all normal RR intervals (SDNN)) after stress exposure are related to a decrease in parasympathetic activity. However, these studies only measured the effects of noise exposure without any prior therapy intervention. But, these studies support the claim that the noise exposure used in Study 2 was stressful.

As mentioned earlier, there were significant increased HR and increased HRV parameters in all conditions at the end of experiment compared to baseline (t4mt1). Despite that, taken all the results from physiological responses together, it was observed that there was a trend that showed decreased SCR (See Figure 4.17, page 144) and lower increased HR (See Figure 4.19, page 148) and lower increased HRV parameters (See Figure 4.21, page 151 for SDNN results, and Figure 4.23, page 153 for HF results), from baseline to post-measurement (t4mt1) in the Paro conditions compared to two other conditions.

Similar trends were observed in changed scores between stress session and at the end of final-measurement (t4mt3). There was a trend that showed SCR, HR, SDNN and HF decreased which almost resulted in a return to baseline in C1:familiar and C2:no_familiar compared to C3:bolster and C4:rest.

These results may suggest Paro conditions help in accelerating faster recovery to stress exposure compared to the other two conditions. These findings are in line with previous studies that mentioned faster recovery in physiological responses after the stress exposure ended (Allen et al. 2002, Amiot & Bastian 2015). A study by Allen et al. (2002) mentioned that the presence of a pet during stressful situation (the stress induction used were cognitive task and physical pressure task) showed the stress reactivity returned to baseline faster than resting alone condition. It is important that the stress responses return quickly to baseline after the stress exposure, in order to reduce the risk of hypertension (Wolff & Frishman 2004).

Effects of conditions on lessening the impact of stress exposure during the stress session

Even though the findings showed no improvement in self-reported and physiological responses in any of conditions at the end of experiment, this section will continue to discuss on the effects of intervention sessions have on stress responses. However, this section will look at the possibility that the earlier session (either Paro or bolster session) might have had a small effect on lessening or minimizing stress responses during Session 2, when compared to Session 1 (t3mt2).

Based on changes in scores from Session 1 to Session 2 (t3mt2), there was a significant increase in self-reported stress and negative affect, and a reduction in self-reported positive affect in all conditions. This suggests that the noise exposure did affect the self-reported stress responses in all conditions.

Then, we look further at the findings on physiological responses, based on changes in scores from Session 1 to Session 2 (t3mt2). In terms of skin conductance response (SCR), the results show that only C4:rest showed a significant increase in SCR, while the other three conditions showed a decreased SCR from Session 1 to Session 2 (t3mt2). The increased SCR in C4:rest may have been the result of the induced stress. This result is in line with (Kumar et al. 2004) and (Park et al. 2018b), where the noise exposures increased skin conductance responses. Then a reduction in SCR (in C1:familiar, C2:no_familiar and C3:bolster) could suggest that the three conditions may have alleviated or minimized the stress responses in terms of SCR.

In the case of heart rate (HR), unexpectedly, HR significantly increased from Session 1 to Session 2 (t3mt2) in all conditions. It was also found that the HR was significantly higher in C1:familiar when compared to C3:bolster. At the same time, there was also a trend that the HR showed an increase in C1:familiar, C2:no_familiar and C4:rest (during Session 2), but a decrease in C3:bolster (See Figure 4.19, page 148). There is a possible explanation for these results, it should be noted that when compared to baseline, there was a trend indicating that the HR was higher in C3:bolster compared to other three conditions, during Session 1. Participants may have felt too relaxed in C1:familiar, C2:no_familiar and C4:rest during Session 1. Then, when participants were exposed to the noise in Session 2, participants may have felt stressed out due to a transition from Session 1 to Session 2, hence the increased HR in C1:familiar, C2:no_familiar and C4:rest compared to C3:bolster. The findings are in line with the opinion of Thoma et al. (2013) that the changes from a relaxing state to stressful induction may have increased the stress responses. These effects probably can be seen in the case of short term relaxing intervention and a sudden stressful event.

In the case of heart rate variability (HRV) parameters, the HRV parameters remained unchanged from Session 1 to Session 2 (t3mt2) in all conditions. This may suggest that the noise exposure did not affect the HRV parameters. These findings are not in line with a study by Sim et al. (2015) and Walker et al. (2016), where the authors found increased HRV parameter during exposure to noise. Walker et al. (2016) suggested that HRV parameters (i.e. high-frequency (HF)) tended to decrease during noise exposure, and in the case of healthy individuals, the HRV normally will return to baseline after noise exposure. But, both studies only measured the effects of noise exposure *per se*. But, in the case of Study 2, participants had an earlier session that was suggested as relaxing (according to lower negative affect scores in all conditions during Session 1). This may suggest that the earlier sessions have had a small effect on HRV parameters and reduced the effects of subsequent stress.

On the basis of the results from HR and HRV, despite increased HR in all conditions, and no changes in HRV parameter across conditions from Session 1 to Session 2 (t3mt2), it could be suggested that participants in C3:bolster and C4:rest were more highly affected by the stress exposure (at Session 2) compared to Paro conditions (C1:familiar and C2:no_familiar).

As mentioned earlier, the increased HR in all conditions may be related to the change from a relaxing state to a stressful state. But, based on an observed trend in HF, the HF was lower in C3:bolster and C4:rest compared to Paro conditions (C1:familiar and C2:no_familiar) (See Figure 4.23, page 153). Therefore, the decreased HF is probably related to decreased parasympathetic activity during stress exposure (as mentioned in an earlier in section 4.4.4, page 114, the HF is used as an indicator of parasympathetic activity (Libby et al. 2012)).

The higher HF in Paro conditions compared to other two conditions may suggest that the earlier interaction with Paro helps to regulate the HF when being exposed to the stress. These results are in line with the opinion of Recio et al. (2016) and Walker et al. (2016), where the authors mentioned that the noise exposure not only increased HR, but at the same time decreased HF.

Effects of familiarisation with Paro on reducing stress responses

As mentioned earlier, *H4(a,b) was not supported*. There was no difference between C1:familiar and C2:no_familiar in all self-reported and physiological responses. But, it was suggested that both Paro conditions had a small effect in helping to minimize the impact of stress induction on physiological responses, but not the psychological responses. This is partially consistent with a previous study by Barker et al. (2010) on the benefits of dog therapy in reducing the stress responses of dog owners and the condition that involved interacting with an unfamiliar dog. The authors found that participants who interacted with familiar and unfamiliar dog had a reduced self-reported and physiological stress response.

Conclusion of Part 2

Overall, the results suggest that none of the four conditions resulted in reduced self-reported responses and HR and HRV parameters as responses to noise exposure. However, there are some observed trends that showed Paro conditions may help in faster recovery compared to other two conditions. This could be due to the high positive moods reported in both Paro sessions, which help in accelerating the faster recovery and minimize the impact of stress. At the same time, it was found that the Paro and bolster conditions may lessen the impact of noise compared to the resting alone condition.

4.13 Limitation and strengths of Study 2

The following are strengths in Study 2 and some limitations that need to be considered:

Stressor:

Firstly, it should be noted that the noise used in Study 1 and 2 was not the same. Study 2 added a new noise of babies crying to the previous noise used in Study 1. Nevertheless, based on the findings in both studies, it was shown that both of the noises were able to induce psychological and physiological stress responses.

Physiological measurements:

Secondly, physiological responses such as skin conductance response (SCR), heart rate (HR), and heart rate variability (HRV) remain a subject that is actively being investigated and that remain controversial in research (Billman 2011). In the case of Paro studies, only a small number of studies have investigated the effects of interacting with Paro using these measures (Mitsui et al. 2001*a*,*b*).

In Study 1, only SCR was used to measure the physiological responses. But, in Study 2, measures of SCR and HRV were used to measure participants' physiological responses. The findings in Study 1 and Study 2 on SCR, showed consistent results indicating arousal when interacting with Paro. The data from HRV in Study 2 adds knowledge to the current literature on physiological responses during the interaction with Paro robot and stress, in terms of measuring the effects of interacting with Paro using different HRV parameters.

4.14 Important consideration for the next study

Based on earlier discussion, and taking together the findings from Part 1 and Part 2, the results of Study 2 indicate that interaction with Paro in Session 1 can induce relaxation and excitement relative to the baseline. However, when compared to baseline and post-measurement, the interaction with Paro did not help in reducing psychological stress responses. Although, there was a trend that suggests an improvement in physiological responses in Paro conditions compared to bolster and resting alone conditions. Even though the results are not as anticipated, it was further suggested that the intervention session conditions might have helped to minimize the effects of stress exposure during the stress session on physiological stress responses.

Therefore, based on the findings of Study 2, the following suggestions are made:

Different intervention formats:

In the next study, Study 3, the plan is to vary the format of interventions, changing order of the stress session and intervention session. This should aid with further understanding the effects that a stress induction and a Paro session on psychological and physiological responses. This will enable, for example, a comparison between an intervention session versus a stress session, and comparing these sessions with the baseline. It will be possible to compare the effects of Paro and stress can be compared without intervention of any effects of previous session. This should help to provide more robust evidence about the on arousal that is related to positive and negative moods in Paro and stress study. Therefore, a further investigation on which types of intervention might have a strong effect on reducing the effects of induced stress on both psychological and physiological responses should be done.

Additionally in Study 3, a further investigation on the effects of Paro on stress will be done by asking participants to do a simple mathematical task during the stressful situation. This will help to find out whether the intervention does really help to reduce or minimize the impact of stress responses, and is even possible that participants will perform better during stress session.

The noise itself has been stressful, as indicated in the findings of Study 1 and Study 2. But, it would be interesting to know if interaction with Paro has a beneficial effect on alleviating the stress responses. In one of short-term studies of Paro by McGlynn et al. (2016), the authors found that there was a reducing trend in stress when performing a stressful task in the presence of Paro. But, in that study, the authors did not measure the impact of Paro on task performance. Additionally, in one of early of studies on effects of noise on task performance, Loeb (1986) reported that a sudden noise had more impact on performance compared to long persistent noise. The authors suggested that the reason was due to the lack of a chance to habituate to the effects of noise, and the finding was in line with study by Szalma & Hancock (2010). But, would the impact of sudden stress be different in Paro intervention study? Therefore, the next study will look at the effects of interacting with Paro on task performance.

Perceived long-term stress:

Furthermore, it was suggested earlier that one of the reasons for less reduction in stress responses in both psychological and physiological responses, could be due to the influence of participants' pre-existing stress. Therefore, there is a need to determine if long term stress is also affecting participants responses to stress and how they interact with Paro.

Therefore, in the next study, prior to the day of the experiment, participants will be asked about their perceived long-term stress.

Photoplethysmography:

Another consideration that should be noted is the use of different physiological measurements to collect the data. For example, in Mitsui et al. (2001*b*) study, the heart rate variability (HRV) was collected using standard lead II of electrocardiogram (ECG) and the sensors were attached on participants' legs. Hence, might explained the difference results between Study 2 and Mitsui et al. (2001*b*).

In Study 2, photoplethysmography (PPG) was used to measure heart rate (HR) and HRV. The PPG has been a reliable method to record HR and HRV signals, and that can be easily attached on the finger (Vescio et al. 2018). Nevertheless, the pulse of PPG are generated based on the contractile of the heart and blood flow from heart to the peripheral tissue. Therefore, the signal could have some time lags which produce errors in the estimation of the pulse period and HRV (Selvaraj et al. 2008). The PPG also has a similar weakness to SCR as the equipment was attached to the finger. The signal has the potential to be affected by movement (Lu et al. 2009). Furthermore, in the study by Lu et al. (2009), who compared the HRV signals derived from PPG and electrocardiogram (ECG), the authors mentioned that the PPG is effective to record heart rate (HR), but not HRV. The HRV signals could be easily affected by the noise produced by movement artefacts. The authors also suggested that the ECG produced more clean HRV signals compared to PPG.

Therefore, in next study, a new measurement of electrocardiogram (ECG) will be used to measure HR and HRV, in order to minimize the effects of artefact movement as suggested by previous studies (Lu et al. 2009, van Erp & Toet 2015).

4.15 Conclusion

Study 2 highlighted the effects of interaction with Paro on subsequent induced stress. The interactions of Paro were compared with bolster and rest sessions. Study 2 has shown promising results of the effects of stroking Paro or bolster on reducing self-reported stress and negative moods during the Session 1. This findings provide a better understanding on effects of stroking active and inactive object on stress responses. Study 2 also provides a better understanding of the effects of novelty and familiarisation in human-robot interaction. It is the first study to have investigated the effects of Paro and a bolster as a preventive short-term environmental stress intervention.

Chapter 5

Study 3: The influence of active Paro and inactive bolster in different stress reducing intervention formats

5.1 Overview of Study 3

5.1.1 Short summary findings of Study 1 and Study 2

In general, the findings in Study 1 and Study 2 revealed that interacting with an active Paro resulted in positive moods. Furthermore, the interaction with Paro most of the time resulted in aroused physiological responses in both studies, and it was suggested that the arousal is related to engagement and relaxation.

Then, in terms of investigating the effects of Paro on reducing stress responses, in Study 1 (on page 99), the findings showed that the physiological stress responses showed a trend towards faster recovery from baseline to post-measurement, and a significantly further reduced from baseline in negative moods at the end of experiment. Nevertheless, in Study 2, there was no stress reduction at the end of experiment when it was compared to baseline for psychological and physiological responses, at least not in the context of using noise as a stressor and looking at the effects of an earlier interaction with Paro on a subsequent stress session. At the same time, Study 2 also compared the effects of interacting with Paro and stroking a bolster and resting alone. Despite no reduction in stress in all conditions, it is worth noting that were some are changes in physiological responses towards faster recovery that were observed when comparing the Paro conditions with other conditions.

From these results (in Study 1 and Study 2), it was suggested that the effectiveness of intervention with Paro may have been influenced by: (i) different intervention formats used, (ii) acute stress and (iii) pre-existing stress, as discussed in the previous chapter (Study 2) on page 160, 170 and 177.

Furthermore, from these results (in Study 1 and Study 2), it is plausible that the positive improvement in moods and physiological responses resulting from the Paro intervention, may also benefit participants' performance. In other words, the effectiveness of the Paro intervention can also be measured by looking at participants' performance of a mathematical task during a stressful situation. At the same time, from these results (in Study 1 and Study 2), it is also possible to determine to what extent the short intervention with Paro may have influenced how participants perceived their existing stress. Study 3 aims to explore how factors such as different intervention formats, pre-existing stress and performance could affect the effectiveness of Paro intervention, or could help to provide more clues about the effects of interacting with Paro.

5.1.2 Aims of Study 3

A further understanding of the effects of interacting with Paro that underlie any effects found in Study 1 and Study 2 is needed. As mentioned earlier, Study 3 aims to explore factors that could influence any effects that occur during the Paro intervention, and its effects on reducing stress responses.

The effects were measured using self-reported measures, physiological responses (skin conductance response (SCR), heart rate (HR) and heart rate variability (HRV): standard of all normal RR intervals (SDNN) and high-frequency (HF)), observed facial expressions and mathematical task performance.

In summary, there are two types of intervention in Study 3: interacting with Paro or bolster (i) before a stress session, or (ii) after a stress session. The stress induction used in Study 3 was a combination of mathematical test and environmental noise (that was used in Study 2). These contrasting interventions should provide more evidence about how interacting with Paro results in reducing the effects of induced stress. Study 3 not only looked at the effects of Paro *per se*, but is also designed to increase an understanding the different effects of interactivity with an active Paro and stroking inactive object (a bolster) on stress reduction. Furthermore, Study 3 also explored whether the short-term intervention session with Paro or bolster would reduce how participants perceived their pre-existing stress.

In short, the following are the new things explored in Study 3:

- Different intervention formats: comparing the effects of intervention session before or after a stress session.
- Can the short intervention session help to improve task performance?
- Can the short intervention session help in reducing how participants' perceived their stress?

5.2 Terms

The terms listed below are defined for the scope of Study 3:

- *Time-point* refers to the time when the physiological measure was taken. There are four time-points in this study where time-point 1 labelled as t1, time-point 2 labelled as t2, time-point 3 labelled as t3 and time-point 4 labelled as t4.

- *After time-point* refers to the time when the participants need to fill in the questionnaires. Participants fill in the questionnaires after each physiological measures time-point. There are four after time-points in this study where after time-point 1 labelled as at1, after time-point 2 labelled as at2, after time-point 3 labelled as at3 and after time-point 4 labelled as at4.

- *Session 1* refers to either an earlier intervention session or earlier stress session at time-point 2.

- *Session 2* refers to either a later intervention session or later stress session at time-point 3, where the participants were asked to listen a noise.

- *Intervention session* refers to either interacting with Paro or stroking a bolster session.

- *Stress session* refers to stress induction session where participants have to listen to a noise while performing a mathematical task.

- *Condition* refers to four conditions in this study where Condition 1 labelled as C1, Condition 2 labelled as C2, Condition 3 labelled as C3, and Condition 4 labelled as C4.

- *Changes in scores* refers to new variables created for changes in scores between time-points. The variables for self-reported responses labelled as at2mat1 (changes in scores between session 1 and pre-measurement), at3mat1 (changes in scores between session 2 and pre-measurement), and at3mat2 (changes in scores between session 2 and session 1). The variables for physiological responses labelled as t2mt1 (changes in scores between session 1 and pre-measurement), t3mt2 (changes in scores between session 2 and session 1), 4mt3 (changes in scores between post-measurement and session 2), and t4mt1 (changes in scores between post-measurement and pre-measurement).

5.3 Rationale of Study 3

5.3.1 Different intervention formats

Based on the results in Study 1 and Study 2, it was found that there were contradictory results in psychological responses at the end of experiment compared to baseline.

In Study 1, decreased self-reported negative affect and stress, and increased self-reported positive affect were found at the end of experiment compared to baseline. But, when the intervention session was changed in Study 2, and participants had an earlier intervention session before the stress session, different effects of intervention session on stress responses were found between Study 1 and Study 2. In Study 2, when comparing between baseline and final measurement, increased self-reported negative affect and stress, and decreased self-reported positive affect were found. These findings suggest that the different intervention formats may have influenced the results in Study 1 and Study 2. It was assumed that the Paro intervention might be more effective as a stress reducer compared to minimizing the effects of subsequent stress. However, the answer to this assumption is still unclear, especially in terms of physiological responses of heart rate (HR) and heart rate variability (HRV), as only Study 2 investigated these measures.

But, in Study 2, the HR and HRV increased from baseline at the end of experiment, and during the Paro intervention session, the HR was low, while, the HRV increased from baseline to the Paro intervention session. Despite that, interestingly, there was a similar trend in skin conductance response (SCR), where the SCR returned to baseline at the end of experiment, and the SCR showed arousal during the Paro intervention session, in both studies.

Therefore, in order to better understand the effects of different intervention formats on psychological and physiological responses further investigation is needed.

Based on the results in Study 1 and Study 2, as mentioned earlier, the results may suggest that Paro intervention more effective to reduce a previous stress (in Study 1) compared to reducing the impact of subsequent stress responses. Thus, in Study 3, *it was anticipated that the later intervention session might have a greater effect in reducing stress compared to the earlier intervention session.*

5.3.2 Effects of noise on performance

There are previous studies in healthy adults that have shown that noise is associated with reduced performance (Smith et al. 2010, Perham et al. 2013, Nasiri et al. 2014).

For example, the study by Smith et al. (2010) found that when participants were exposed to noise while doing mental mathematical task, participants were disturbed and their performance was worse than when doing the task in a quiet session. In a different study by Perham et al. (2013), the authors also investigated the impact of noise on performance. The authors found that the noise not only affected participants' performance in math but also affected the memory ability to recall a series of numbers when compared to a quiet session. Both studies used office noise i.e. crowd conversation, and the sound of a phone ringing and people walking, in order to create mild stress while performing the task.

Then, in a review by Szalma & Hancock (2011) reported that participants were likely to have been more affected by a sudden noise than by listening to noise continuously for a longer time. The sudden noise was an immediate sound that came after a resting session. The sudden noise also tended to affect performance more than continuous noise. The authors suggested that when being exposed to continuous noise, participants may develop habituation to the noise, hence reducing their stress. Then, a study by Nasiri et al. (2014) mentioned that increased noise levels were also related to performance while performing a mathematical task. Participants tended to take a longer time to complete the mathematical task when they had been exposed to higher noise levels compared to lower noise levels. However, it should be noted that, these studies mentioned earlier focused on doing a mental mathematical task, but in Study 3, participants were asked to do mathematical task on paper. The reason for allowing participants to use paper was to ensure that the stress session would not induce too much stress on participants.

These past research showed evidence of the detrimental effects of noise on adults task performance. Therefore, in general, *it was anticipated in Study 3 that the noise stress session would not only influence the participants' stress responses, but would also reduce participants' performance while doing the mathematical task.*

5.3.3 Effects of therapy intervention on performance

Based on the review in the previous section, past research confirms that stress can affect performance.

However, in terms of Paro intervention and performance, there is a little research on how short-term interaction with Paro will help in reducing stress and improving performance. In studies by Kawaguchi et al. (2012) and Wada et al. (2005b), the authors found that interacting with Paro results in changes in activation in some parts of the brain. The authors suggested that the brain activation can lead to improvements in cognitive abilities which help to improve memory and the ability to solve problems (Kawaguchi et al. 2012). Another study of Paro by McGlynn et al. (2016) investigated the stress-reducing effects of Paro on induced stress that was related to a cognitive task. The study found that there was a decreased trend in stress levels when participants were asked to perform a stressful task compared to when no Paro was present during the stress task.

However, both studies by Kawaguchi et al. (2012) and Wada et al. (2005b) did not focus on investigating the impact that interacting with Paro has on task performance. In a recent on-line article in 2018, Shibata suggested that there is the potential that Paro can be used in aerospace to help to reduce astronaut stress (Holland 2018). It was further mentioned that the reduced stress may contribute to prevent the impact of stress on performance that may leads to human error[5] (Holland 2018). However, no exploratory study yet has been done.

[5] In industrial engineering, human error was defined as a disruption occurs during operational task that potentially damaging the equipment, as results from a poor performance to a designed task (Dhillon & Liu 2006).

Despite that, the studies by Kawaguchi et al. (2012), Wada et al. (2005b) and McGlynn et al. (2016) provide an insight into the influence of Paro in improving brain activity, and minimizing the effects of stress from the cognitive task. Nevertheless, there are also past studies that found an effect of animal therapy on arithmetic performance that focused on the effects of short-term interaction with animal therapy during (Allen et al. 2002) or before (Trammell 2017) stress manipulation.

In a study by Allen et al. (2002), participants were asked to perform a stressful task while accompanied by a dog, and it was found that the presence of the dog helped to minimize the impact of the task on participants' cardiovascular reactivity. The reduced responses to stress may have suggested that the dog helped participants to perform a task. However, in a study by Allen et al. (2002), the authors did not measure the impact of dog session on task performance. Later, a study by Trammell (2017) found that participants who had a short session of dog therapy before an upcoming examination day, performed better than participants who did not. It was suggested that participants performed better in the exam as their levels of stress was reduced after the dog therapy.

Based on the positive benefits of pet or Paro intervention in past research, Study 3 aims to explore the effects of Paro intervention on participants' task performance during the stress session. *It was anticipated that interacting with Paro prior to a stress session may help to reduce participants' stress and helps the participants to perform better than participants who do not have a session with Paro before the stress session.*

5.3.4 Perceived stress

Perceived stress is defined as how an individual determines that a situation that happened in their life was stressful and how frequently they feel like that (Cohen et al. 1994, Stubbs et al. 2017). For example, the perceived stress can be measured using Perceived Stress Scale (PSS) questionnaire, and one of the questions is "*In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?*" The Likert-Scale range from 0

(Never) to 5 (Very often) (Cohen et al. 1994).

Past studies of Paro have focused on the use of Paro in long-term interaction and findings found a positive improvement in elderly people with dementia with reduce anxiety, stress, agitation and increase social interaction with other people, as reviewed in Chapter 2.3, page 16. However, there is no study of Paro that focuses on whether the Paro intervention can reduce the perceived stress responses, or whether the perceived stress may have influenced the ways of interaction with Paro.

In terms of social support and perceived stress, a study by Lakey & Cohen (2000) found that companionship (such as interacting with pet) may acts as a moderator to reduce the perceived stress. It was also found by Civitci (2015) that there was a relation between negative affect and perceived stress. Civitci (2015) suggested that high perceived stress in individuals might reduce the effectiveness of the social support, and that the effects were also influenced by high negative affect experienced by an individual.

As mentioned earlier that there are no past studies of Paro that have investigated on the relation between perceived stress and Paro intervention. Nevertheless, some literature mentions reasons for elderly people to reject the use of Paro robot (Klein & Cook 2012, Wagemaker et al. 2017). One of the reasons can be that the Paro reminds them about their hard times in past, hence reducing their interest to interact with Paro (Klein & Cook 2012). In addition, a recent on-line article by Randall (2017) also mentioned that one of the elderly people with dementia in nursing home suddenly shouted angrily towards Paro after not talking for a very long time. These negative reactions toward Paro may have been related to how the elderly perceived their stress, hence, it may have reduced the effectiveness of any Paro intervention. However, even though negative reactions toward Paro leaves staffs baffled, it is still plausible that the presence of Paro may have a positive potential to encourage people with dementia to improve their daily communication skills (Randall 2017).

Therefore, in case of Study 3, *it was anticipated that participants that have an intervention* *with Paro would show lower perceived stress at the end of experiment compare to baseline.* This anticipation made based on the low self-reported negative affect and high self-reported positive affect found during Paro intervention in Study 1 and Study 2.

5.3.5 Summary method of Study 3

In Study 3, an earlier interaction with Paro or bolster session or an earlier stress induction are labelled as Session 1 (S1). A session with Paro or bolster after the stress induction or a later stress induction after interaction session are labelled as Session 2 (S2). At the same time, all Paro or bolster sessions will be called the experimental conditions or intervention session, while the noise and mathematical session, will be called a stress session.

Repeated measurement:

Study 3 had four time points (t1 to t4) of 5 minutes' measurement. A pre-measurement (t1) was taken for 5 minutes (at t1). Then this was followed by 5 minutes' session with either Paro or bolster session, or a stress session (S1) (at t2), then 5 minutes' of either a Paro or bolster session, or a stress session (S2) (at t3) and a 5 minutes' post-measurement (at t4). A detailed experimental procedure for Study 3 is described in section 5.7, page 198.

Four conditions:

Study 3 also had four conditions and participants were randomly allocated to one of the four conditions, as follows:

- In condition 1 (C1), participants had an earlier session with Paro before they were subjected to a session designed to induce mild stress. The mild stress inducements used in Study 3 were noise and a mathematical task.
- In condition 2 (C2), participants were first subjected to a session designed to induce mild stress, and were then given a session with Paro.

- In condition 3 (C3), participants did not have a chance to interact with Paro, but were given a soft bolster to stroke. After the session with the bolster, participants had a stress session.
- In condition 4 (C4), participants were assigned to a stress session that followed by an intervention session (stroking a bolster).

5.4 Hypotheses

Based on overall aims of Study 3 on page 182, the hypotheses of Study 3 are as follows:

Hypothesis 1 (H1):
H1 is about whether different intervention formats may have a different impact on induced stress based on changes in scores between baseline and final-measurement (at3mat1 or t4mt1).

As mentioned in section 5.3.1, it was suggested that the stress reducing effects of interacting with Paro may have been more effective in reducing the effects of an earlier stress induction compared to the effects of a later stress induction.

H1(i): It was predicted that a session with a Paro (C1 and C2) compared to a bolster (C3 and C4) condition would result in:

H1(i)(a) a reduction in self-reported stress and negative affect, and an increase in self-reported positive affect. (at3mat1)

H1(i)(b) a reduction in SCR, HR, and improvement in SDNN and HF. (t4mt1)

H1(ii): It was predicted that a later session with Paro in C2 compared to an earlier session with Paro in C1 would result in:

H1(ii)(a) a reduction in self-reported stress and negative affect, and an increase in self-reported positive affect. (at3mat1)

H1(ii)(b) a reduction in SCR, HR, and improvement in SDNN and HF. (t4mt1)

Hypothesis 2 (H2):

H2 is about how the different intervention formats may influence participants' responses during the intervention session.

As mentioned on page 185, it was suggested that the stress reducing effects of interacting with Paro may have been more effective in reducing an earlier stress induction compared to a later stress induction.

It was predicted that an intervention with a Paro (in C1 and C2) compared to a bolster session (in C3 and C4) would result in:

H2(a) low self-reported stress, negative affect, negative emotions, and high self-reported positive affect and positive emotions

H2(b) low HR, and high SCR, SDNN and HF

H2(c) higher frequency of observed positive and lower negative facial expressions

Hypothesis 3 (H3):

H3 is about how the different intervention formats may influence participants' responses during the stress session.

H3(i): It was predicted that participants in the later stress session (in C1 and C3) compared to earlier stress session (in C2 and C4) would have:

H3(i)(a) low self-reported stress, negative affect, and higher self-reported positive affect

H3(i)(b) low HR and SCR, and improve in SDNN and HF

H3 (ii): Then, it was further predicted that participants in the later stress session in C1 (an earlier session with Paro) compared to participants in the later stress session in C3 (an earlier session with bolster) would have:

H3(ii)(a) low self-reported stress, negative affect, and higher self-reported positive affect

H3(ii)(b) low HR and SCR, and improve in SDNN and HF.

Hypothesis 4 (H4):

H4 is about how the different intervention formats may influence participants' performance when doing a mathematical task during the stress session.

It was anticipated that participants that have a session with Paro before the stress session (in C1) would perform better than participants in bolster session (C3), and compared to the other two conditions (C2 and C4), that have an intervention session after the stress session.

Hypothesis 5 (H5):

H5 is about how the different intervention formats may influence participants' perceived stress, based on a comparison between baseline (t1) and post-measurement (t4).

It was predicted that a session with a Paro (C1 and C2) compared to bolster (C3 and C4) condition would result in lower self-reported perceived stress scores (t1 and t4).

5.5 Methods

5.5.1 Participants

Participants were recruited through the University of Sheffield volunteering distribution-emailing list. A small announcement was also placed on Student Services information Desk (SSiD). Participation was voluntary and all information was treated as confidential. Participants were told that they were free to leave at any time during experiment. Participants were rewarded with £5 for their time. Exclusion criteria in Study 3 were similar to Study 1 and 2, as described in previous chapter, see Chapter 3, section 3.7.1, page 72.

104 participants were recruited to take part in the study (66 females, 42 males). None of the participants had any prior experience playing with Paro. Participants were aged between 18 to 35, ($M = 23.34, SD = 3.790$). Half of participants were White, 54 (51.92%), the remainder of the participants were Asian, 30 (28.85%), Mixed, 9 (8.65%), Black, 7 (6.73%) and other 4 (3.85%). Most participants were studying courses in Science and Engineering, 45 (43.75%), other participants were from Medicine, 22 (21.15%), Social Science, 18 (17.31%), and Arts and humanities, 19 (18.27%).

Table 5.1 shows summary of age and gender for each condition. A one-way between-subjects ANOVA (gender as dependent variable and condition as between-subjects factor) found there was no significant difference in gender between conditions, $F(3,100) = 0.051, p = 0.985, r = 0.002$. A one-way between-subjects ANOVA test also (age as dependent variable and condition as between-subjects factor) found there was no significant difference in age between conditions, $F(3,100) = 1.495, p = 0.221, r = 0.043$.

Table 5.1: Report of age and gender by condition for Study 3

Condition	*N*	Age	Gender
1	26	M = 23.39; SD = 4.129	M = 10; F = 16
2	26	M = 24.54; SD = 3.636	M = 10; F = 16
3	26	M = 22.38; SD = 3.112	M = 11; F = 15
4	26	M = 23.04; SD = 4.084	M = 11; F = 15
Total	104	M = 23.34; SD = 3.790	M = 42; F = 62

5.6 Measurement and materials

5.6.1 Mathematical task as a stressor

The mathematical task of consisted questions from General Certificate of Secondary Education (GCSE) previous examination papers. The level of the mathematical task is easy. Participants were not allowed to use a calculator but were given a paper and pencil to do the calculation. There are 9 questions that were asked in a 5 minute period, see section A.10, page 329. Participants were given 30 seconds to answer each question, see example on Figure 5.1. All questions and the 30 seconds time count down for each question was displayed on the screen.

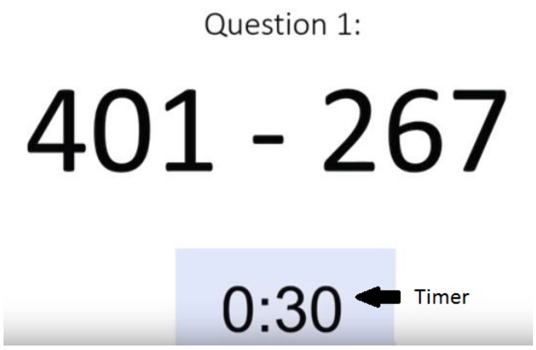


Figure 5.1: Mathematical questions

5.6.2 Physiological measurement

In Study 3, two physiological sensors were used to record participants' responses: (i) skin conductance and (ii) electrocardiography (ECG), to measure heart rate and heart rate variability.

5.6.3 Physiological data acquisition

A detailed description about skin conductance has been discussed earlier in this thesis, see Chapter 3, section 3.8.5, page 75. However, the two physiological signals were recorded wireless using BIOPAC MP160, to record the real-time physiological responses data, see Figure 5.2. All sensors have a sampling rate of 200 Hz and the raw signal was 0.05-35 Hz bandpass-filtered (Hoffmann et al. 1996, Lu et al. 2009). As mentioned earlier in the literature in Chapter 2 on page 47, heart rate and heart rate variability data was acquired through electrocardiogram (ECG) signal. Three pre-gelled disposable electrodes were used to obtain the signal and two electrodes were placed on left (black cable) and right clavicle (white cable) and just above the left waist (red cable).

Figure 5.2: BIOPAC MP160

5.6.4 Video recording

Videos were recorded during the 5 minutes that participants were exposed to or interacted with Paro and bolster. The videos were recorded using web camera (Logitech C920 HD Pro 1080p).

5.6.5 Perceived stress scale

Perceived stress scale (PSS-10) asked about participants' pre-existing stress within one-month. Participants rated their responses using five point Likert-Scale (0 to 4) from *Never, Almost never, Sometimes, Fairly often, and Very often* (Cohen & Janicki-Deverts 2012). Participants with scores below 13 or 19 considered as average or have low stress, and participants with high scores which is above 19, perceived their stress as high Cohen & Janicki-Deverts (2012).

5.6.6 Mathematical anxiety

The modified mathematical anxiety scale (MAS) adopted from study by Mahmood & Khatoon (2011) is a 14-item questionnaire that measure the mathematical anxiety of students. Questionnaires are rated on 5-point Likert-Scale. Total scores can range from 14 to 70. Higher scores reflect greater mathematical anxiety. MAS has split-half reliability of 0.89 and Cronbach's $\alpha = 0.87$.

5.6.7 Noise annoyance

Additional questions about the participants' self-reported experience of the noise were asked. The questions were adapted from previous study by Fields (1998), Gidlof-Gunnarsson et al. (2007), Alvarsson et al. (2010) who used noise as a stressor in their studies. In these past studies, participants were asked about the pleasantness or unpleasantness of the noise, for example, *Did the noise annoy you?*. The noise annoyance is defined as "*a feeling of displeasure caused by noise and the feelings of discomfort to the noise because it interferes with the thoughts or activity*" (Lindvall & Radford 1973, Passchier-Vermeer & Passchier 2000, Sung et al. 2017).

In Study 3, participants were asked to rate how they felt about the noise ("*Did you find the noise annoying?*"). Participants answered on 7-point of Likert-Scale ranges from 1 = "Not at all" to 7 = "Extremely').

5.6.8 Other materials

Other materials and measurements including Paro, bolster (see section 4.7.1, page 118), Positive and Negative Affect Schedule (PANAS) and levels of stress questionnaire, were identical to those used in Study 1 and 2. All detailed descriptions about Paro, PANAS, self-reported stress and emotions have been discussed earlier in this thesis, see Chapter 3, section 3.8.1 on page 73, section 3.8.3 on page 74, and section 3.8.4 on page 74, and 4.7.6 on page 121.

5.7 Experimental procedure

As mentioned in previous section 5.5.1, the study was advertised on-line and through email. Ethical approval was obtained from University of Sheffield, Department of Computer Science Research Ethics Review Committees (Reference number: 011158). Interested participants filled in an on-line survey through Qualtric (an on-line survey tool). The initial survey asked about their demographic and perceived stress scale (PSS-10), see section A.13, page 331. At the end of the survey, participants were asked to set an appointment for the experiment.

The day before the experiment day, a reminder was sent to participant to remind about the scheduled appointment. Experiments were conducted in a listening room at Department of Computer Science, University of Sheffield. The room contains a chair and a small table, a keyboard, a mouse, a video, a headphone and a camera. A monitor was located on the outside of a see-through barrier, see Figure 5.3 and Figure 3.1 on page 73 for an overview of the listening room.

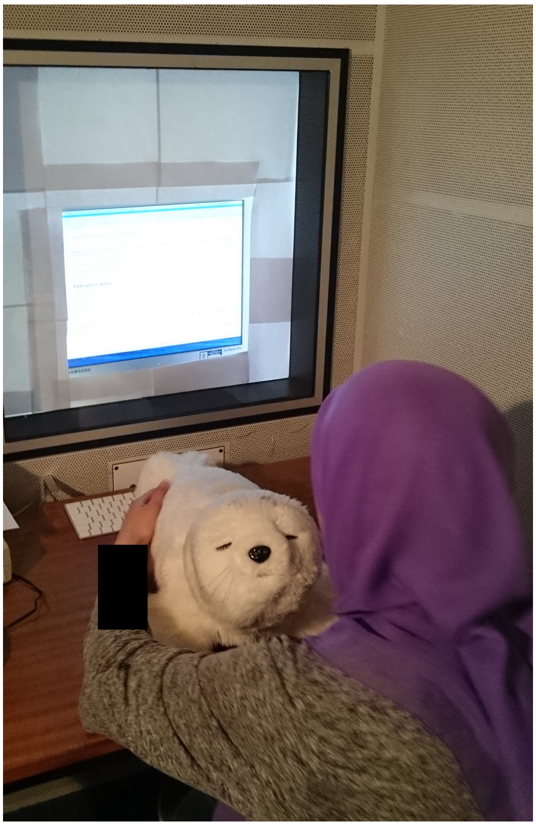


Figure 5.3: Listening booth

On the experiment day, participants were asked to read participation information sheet. The participant information sheet included a description of the purpose of the study and an overview of the experimental procedure, see section 325, page 325. Then participants provided their consent to participate in this experiment on the computer. The experiment began with experimenter giving an orally instruction about the experimental procedure to participants. The experimental procedure varied depending on the condition that participant was assigned to.

The experiment consisted of four conditions: (C1) Paro followed by stress induction (mathematical task and noise), (C2) stress induction (mathematical task and noise) followed by Paro, (C3) bolster followed by stress induction (mathematical task and noise), (C4) stress induction (mathematical task and noise) followed by bolster. *The four conditions were labelled as follow:* C1:Paro_stress, C2:Stress_Paro, C3:Bolster_stress, C4:Stress_bolster.

Participants were assigned randomly to one of four conditions. In Study 3, time-point 1 (t1) was called *Session 1* and time-point 2 (t2) was called *Session 2*. In condition 1 (C1:Paro_stress), participants were asked to play with Paro for 5 minutes in Session 1. Then they were asked to perform a mathematical task and listen to a noise for 5 minutes in Session 2. For condition 2 (C2:Stress_Paro), participants were asked to perform a mathematical task while listening to a noise for 5 minutes in Session 1. Then they were asked to play with the Paro for 5 minutes in Session 2. In condition 3 (C3:Bolster_stress), participants did not have a session with the robot but were asked to stroke a bolster in Session 1 (S1) and then listen to a noise while answering mathematical task in Session 2 (S2). In condition 4 (C4:Stress_bolster), participants were asked to answer the mathematical task in and listening to a noise in Session 1 then stroking a bolster in Session 2.

Physiological sensors were attached to participants' fingertips (for skin conductance (SC)), and clavicle and waist (for heart rate (HR) and heart rate variability (HRV)). Then participants were asked to sit comfortably and asked to wear headphones.

Participants were asked to follow all the instruction on-screen. The experiment started after the experimenter left the room and closed the door and participants clicked the 'Next' button on-screen.

The experiment lasted around 30 to 45 minutes. The physiological measures were taken for 5 minutes at four time points (t1 - t4). The physiological measures were recorded at pre-measurement (t1), and during Session 1 (t2), Session 2 (t3) and post-measurement (t4).

At each measurement time-point, participants were asked to "Please wait until the experimenter tells you to click 'Next' button" on-screen. Then the experimenter showed a 'Next' card in front of the monitor. This is to make sure all the physiological measures were recorded at each 5 minute.

After each measurement, at after time-point 1 (at1), after time-point 2 (at2) and after time-point 3 (at3), participants completed questionnaires: stress levels and positive and negative affects (PANAS).

The experimenter pointed out that there was a video camera and the video would be recorded when they interacted with Paro or stroked the bolster. The video camera would only record their faces during this time. They were allowed to refuse to be recorded. However, in Study 3, none of the participants refused to be in the video.

After stress induction (mathematical task and noise), the experimenter would immediately enter the room to collect the paper and pencil. This was to prevent participants writing, looking, or playing with the paper and pencil. Participants were not allowed to talk with the experimenter.

After the last measures (t4), participants were asked to complete a few questionnaires. The questionnaires were asking about their perceived-stress scale (PSS-10), noise experience and feelings while stroking the bolster or interacting with the Paro. After they finished answering all questionnaires, all the sensors attached to participants were removed. Before leaving the room, participants were given time to ask further questions about this study. Participants were compensated for their participation with £5.

The procedure of experiment conducted in Study 3 can be seen in Figure 5.4.

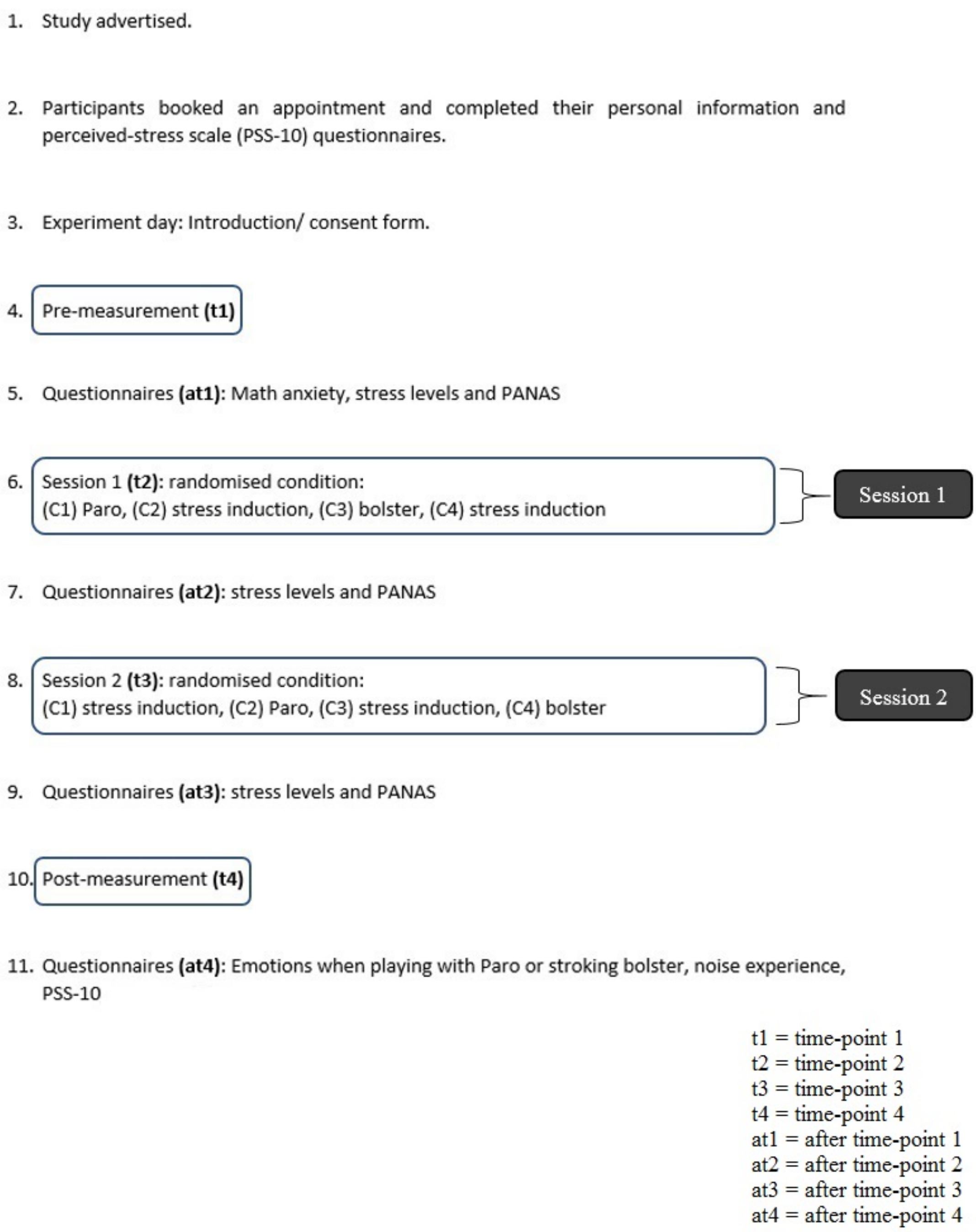


Figure 5.4: Study 3 experimental procedure

5.8 Data analysis

All data collected were analysed using the same way as in Study 2, refer section 4.9 on page 123.

In summary for data analysis flow, at first the data were checked for normality. Then the data were computed using descriptive statistics for demographic variables. One-way analysis of variance used to determine conditions differences in baseline of each measure. The one-way ANOVA also been used to analysed effect of conditions on self-reported emotions, observed facial expressions, mathematical anxiety and mathematical task scores. Repeated measures were analysed using a mixed between-within subjects' ANOVA to determine the overall effect of the independent variable (i.e. self-reported stress, PANAS, perceived stress, physiological responses) on the conditions.

However, there were some new analyses done in order to answer the new research questions in Study 3. Further analyses were used to determine relationship between dependent variable(s) and conditions for perceives stress, noise experience and mathematical scores using McNemar or chi-square.

Heart rate variability data: Data collected from heart rate variability were pre-processing off-line using AcqKnowledge 4.4 software. The low frequency data were cut off at 1 Hz and high frequency were cut off at 35 Hz. Then the ECG data were visually inspected and artifacts were corrected or if necessary removed manually (Dotsinsky 2007, Krypotos et al. 2011). The time domain index of heart rate variability (HRV) used was the standard deviation of normal to normal R-R intervals (SDNN). The time-frequency used was high frequency (Wang & Huang 2012).

Heart rate data: The heart rate data collected from ECG were pre-processed using Kubios 2.0 software (Tarvainen et al. 2014). The signals were corrected using the default option "medium artifact correction" in Kubios software. Additionally, to remove disturbing low frequency baseline trend components, a Lambda value of 500 was set to the "smoothness prior method" option (Luque-Casado et al. 2013). Then the data were manually inspected and corrected for any artefact on or removal (Krypotos et al. 2011).

5.9 Results

5.9.1 Baseline of each measure

There were no significant differences between conditions at pre-measurement (at1) in any of the self-reported and physiological responses. This finding indicates that participants in all conditions had similar levels of stress, Positive and Negative Affect Schedule (PANAS), and physiological arousal at the beginning of the study, all $p > 0.05$.

Baseline mathematical anxiety was also measured to determine all conditions had a similar levels of mathematical anxiety. There was no significant difference between conditions on mathematical anxiety scores at pre-measurement (at1), $F(3,100) = 1.877, p = 0.138, r = 0.053$. Further analysis using Chi-square to analyse the relationship between categories of mathematical anxiety at pre-measurement (low, high) and conditions. The mathematical anxiety scores for each time-point were split into two categories, *Low mathematical anxiety* and *High mathematical anxiety*. Participants with total scores that higher than 32 were considered as high level of mathematical anxiety (Mahmood & Khatoon 2011, Johnston-Wilder et al. 2014). There was no significant difference between categories of mathematical anxiety at pre-measurement (low, high) and conditions, $X^2(3, N = 104) = 2.213, p = 0.529$. But as can be seen in cross-tabulation Table 5.2, most participants rated high in mathematical anxiety for all conditions at baseline. The result indicates that participants in all conditions had a similar level of mathematical anxiety.

Table 5.2: Two-way table of counts and percentages for the relationship between categories of mathematical anxiety at pre-measurement (low, high) and conditions

Mathematical anxiety categories		Condition 1	Condition 2	Condition 3	Condition 4
Low	Count	1	4	2	3
	% within Condition	3.8%	15.4%	7.7%	11.5%
High	Count	25	22	24	23
	% within Condition	96.2%	84.6%	92.3%	88.5%

5.9.2 Effects of conditions on self-reported responses

Self-reported stress

A mixed between-within subjects analysis of variance showed there was a significant effect of time on stress levels over three time points, $F(2,200) = 20.01, p < 0.0001, r = 0.167$, and no significant effect of conditions, $F(3,100) = 0.336, p = 0.799, r = 0.010$. There was a significant interaction between condition and time on stress levels, $F(6,200) = 53.45, p < 0.0001, r = 0.616$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. The stress level scores showed a significant increase in C2:Stress_Paro and C4:Stress_bolster, and, decrease in C1:Paro_stress and C3:Bolster_stress from pre-measurement (at1) to Session 1 (at2) , $p < 0.0001$. The stress level scores then showed no significant difference from Session 1 (at2) to Session 2 (at3), $p = 1.000$. However, there was a significant increase in C1:Paro_stress and C3:Bolster_stress and, decrease in C2:Stress_Paro and C4:Stress_bolster between pre-measurement (at1) and Session 2 (at3), $p < 0.0001$. (Refer to Figure 5.5).

A further analysis on score differences was carried out in next paragraph, in order to check for significant difference between conditions.

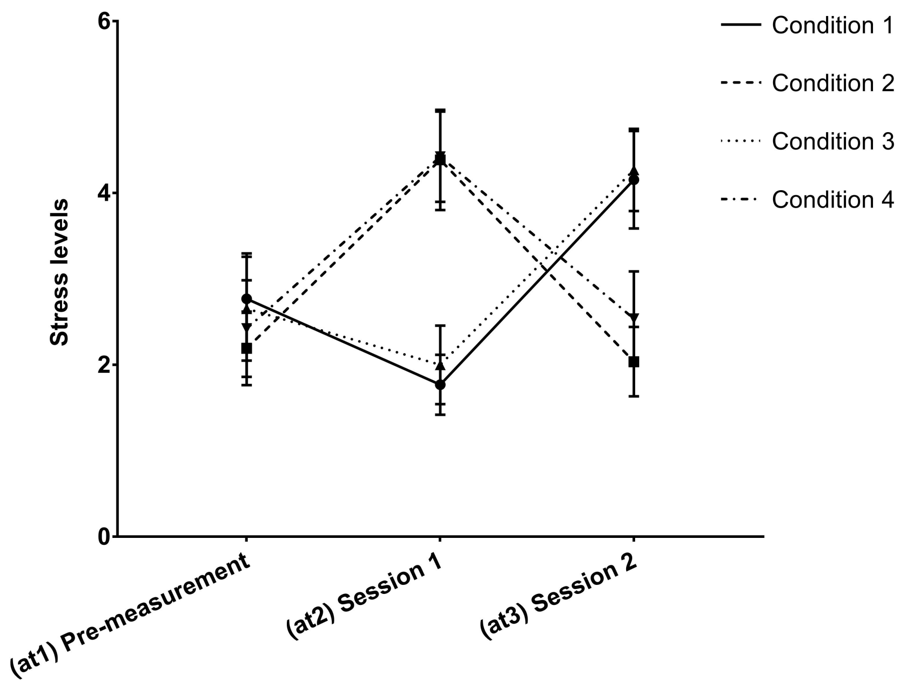


Figure 5.5: The mean with 95% confidence interval for self-reported ratings on levels of stress to each condition.

New variables: New variables from paired time points were created in order to check if there is difference between conditions in terms of changes in stress level scores. The new variables are $SL\_at2mat1$ = Session 1 minus pre-measurement, $SL\_at3mat2$ = Session 2 minus Session 1 and $SL\_at3mat1$ = Session 2 minus pre-measurement. The graph for the results is shown in Figure 5.11 on page 215.

There was a significant difference between conditions for $SL\_at2mat1$, $F(3,100) = 41.54, p < 0.0001, r = 0.555$. The changes in self-reported stress between pre-measurement and Session 1 showed a greater decrease in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress also showed a greater decrease compared to C4:Stress_bolster, $p < 0.0001$. However, C2:Stress_Paro showed a greater increase in self-reported stress compared to C3:Bolster_stress, $p < 0.0001$. There was no significant difference between C3:Bolster_stress and C4:Stress_bolster in self-reported stress. There is evidence here that the stress induction at Session 1 in C2:Stress_Paro and C4:Stress_bolster resulted in a greater level of reported stress.

There was a significant difference between conditions for $SL\_at3mat2$, $F(3,100) = 103.74, p < 0.0001, r = 0.757$. The changes in self-reported stress between Session 1 and Session 2 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress also showed a greater increase compared to C4:Stress_bolster, $p < 0.0001$. C1:Paro_stress and C3:Bolster_stress and, C2:Stress_Paro and C4:Stress_bolster showed no significant differences in self-reported stress, $p = 1.000$.

There was a significant difference between conditions for $SL\_at3mat1$, $F(3,100) = 13.31, p < 0.0001, r = 0.285$. The changes in self-reported stress between pre-measurement and Session 2 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress and C4:Stress_bolster showed a significant difference but both conditions showed an increase in stress levels, $p < 0.0001$. C1:Paro_stress and C3:Bolster_stress and, C2:Stress_Paro and C4:Stress_bolster showed no significant differences in self-reported stress, $p = 1.000$.

Paro or bolster session A one-way between-subjects ANOVA showed no significant difference between conditions in self-reported stress at Paro and bolster sessions, $F(3,100) = 2.235, p = 0.089, r = 0.063$. Although no significant difference at Paro or bolster sessions but as can be seen in Figure 5.6, the stress levels in C1:Paro_stress were lower compared to other conditions. However, C4:Stress_bolster showed higher levels of stress compared to other conditions.

Stress session A one-way between-subjects ANOVA showed no significant difference between conditions in self-reported stress at stress session, $F(3,100) = 0.215, p = 0.886, r = 0.006$. Although no significant difference at stress session, but as can be seen in Figure 5.6, the stress levels in C1:Paro_stress were lower compared to other conditions. However, C4:Stress_bolster showed higher levels of stress compared to other conditions.

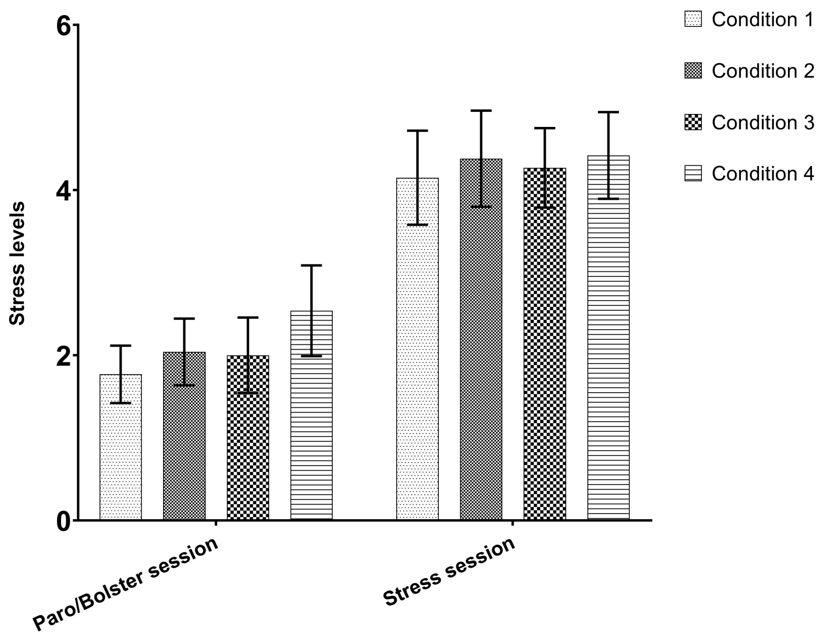


Figure 5.6: The mean with 95% confidence interval for scores on levels of stress at Paro and bolster and stress session to each condition.

Positive affect in PANAS

A mixed between-within subjects analysis of variance showed there was a significant effect of time on positive affect, $F(2,200) = 6.138, p = 0.003, r = 0.058$, and a significant effect of conditions, $F(3,100) = 6.270, p = 0.001, r = 0.158$. There was a significant interaction between condition and time on positive affect, $F(6,200) = 3.296, p = 0.004, r = 0.090$.

Pairwise comparisons between time points were done using Bonferroni correction test. The positive affect scores showed no significant difference between pre-measurement (t1) and Session 1 (t2), $p = 1.000$. The positive affect scores then showed a significant decrease from the Session 1 (at2) to the Session 2 (at3), $p = 0.014$. However, there was a significant decrease from pre-measurement (at1) and Session 2 (at3), $p = 0.009$. (Refer to Figure 5.7).

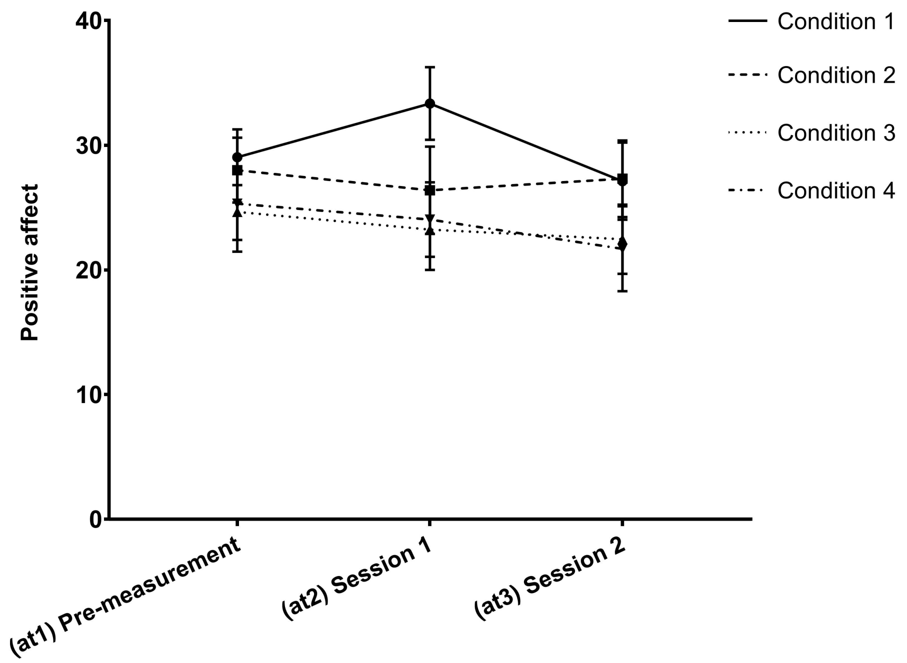


Figure 5.7: The mean with 95% confidence interval for positive affect in PANAS for each condition.

New variables: New variables from paired time points were created in order to check if there is difference between conditions in terms of changes in positive affect scores. The new variables are $PA\_at2mat1$ = Session 1 minus pre-measurement, $PA\_at3mat2$ = Session 2 minus Session 1 and $PA\_at3mat1$ = Session 2 minus pre-measurement. The graph for the results is shown in Figure 5.12 on page 216.

There was a significant difference between conditions for $PA\_at2mat1$, $F(3,100) = 4.752, p = 0.004, r = 0.125$. The changes in PA scores between pre-measurement and Session 1 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p = 0.012$, C3:Bolster_stress, $p = 0.016$, and C4:Stress_bolster, $p = 0.021$. There was no significant difference between C2:Stress_Paro and C3:Bolster_stress, between C2:Stress_Paro and C4:Stress_bolster and between C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$.

There was a significant difference between conditions for $PA\_at3mat2$, $F(3,100) = 4.410, p = 0.006, r = 0.117$. The changes in PA scores between Session 1 and Session 2 showed a greater decrease in C1:Paro_stress compared to C2:Stress_Paro, $p = 0.005$. However, the C2:Stress_Paro showed an increase in changes in PA scores. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, $p = 0.056$, between C1:Paro_stress and C4:Stress_bolster, $p = 0.372$, between C2:Stress_Paro and C3:Bolster_stress, $p = 1.000$, between C2:Stress_Paro and C4:Stress_bolster, $p = 0.693$, and between C3:Bolster_stress and C4:Stress_bolster, $p = 1.000$.

Similar result was found at $PA\_at2mat1$ and $PA\_at3mat2$ for PA, interacting with Paro induce positive affect. Regardless when the Paro session either in Session 1 or Session 2, the Paro helps in inducing positive affect.

There was no significant difference between conditions for $PA\_at3mat1$, $F(3,100) = 0.749, p = 0.525, r = 0.022$. All conditions showed a decrease in positive affect. But, as can be seen in Figure 5.12, the condition with Paro (C1:Paro_stress and C2:Stress_Paro) showed a lower decrease in positive affect compared to C3:Bolster_stress and C4:Stress_bolster.

Paro or bolster session A one-way between-subjects ANOVA showed there was a significant difference between conditions at Paro and bolster session, $F(3,100) = 11.50, p < 0.0001, r = 0.256$. There was significant difference between C1:Paro_stress compared to C2:Stress_Paro, $p = 0.039$, C3:Bolster_stress, $p < 0.0001$, and C4:Stress_bolster, $p < 0.0001$. As shown in Figure 5.8, the positive affect scores were higher in C1:Paro_stress compare to C2:Stress_Paro, C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C2:Stress_Paro and C3:Bolster_stress, $p = 0.380$, between C2:Stress_Paro and C4:Stress_bolster, $p = 0.067$, and between C3:Bolster_stress and C4:Stress_bolster, $p = 1.000$.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session, $F(3,100) = 2.025, p = 0.115, r = 0.057$. Although no significant difference but as can be seen in Figure 5.8, the positive affect in C3:Bolster_stress and C4:Stress_bolster was lower compared to Paro conditions (C1:Paro_stress and C2:Stress_Paro).

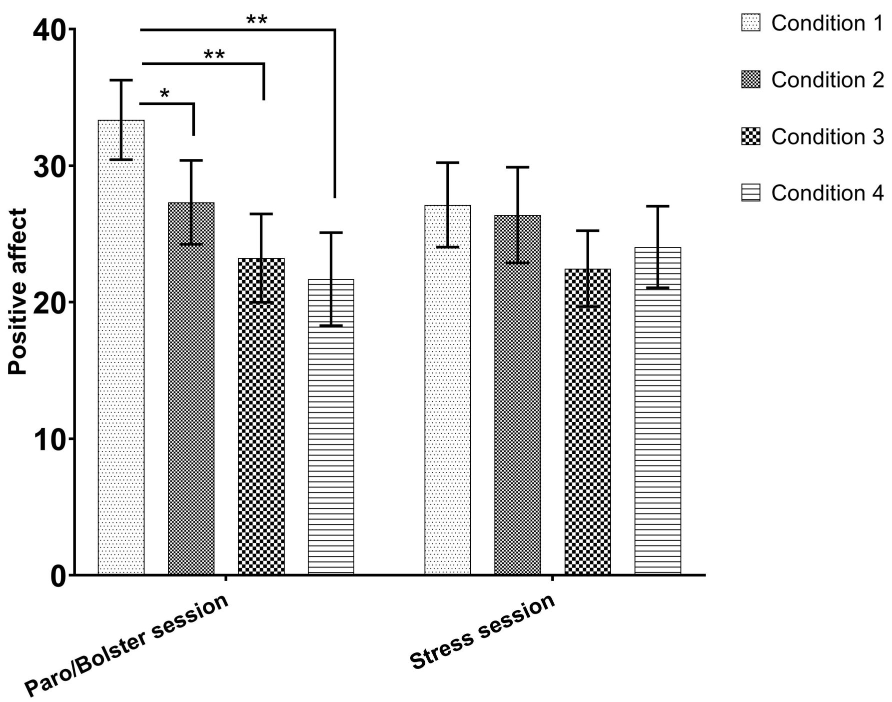


Figure 5.8: The mean with 95% confidence interval for positive affect at Paro and bolster and stress session to each condition. $*p < 0.05, **p < 0.001$

Negative affect in PANAS

Two outliers were removed from the analysis. A mixed between-within subjects analysis of variance showed there was a significant effect of time on negative affect, $F(1.842, 180.49) = 6.137, p = 0.003, r = 0.059$, and no significant effect of conditions, $F(3,98) = 0.240, p = 0.869, r = 0.007$. There was a significant interaction between condition and time on negative affect, $F(5.525, 180.49) = 33.61, p < 0.0001, r = 0.507$.

Pairwise comparisons between time points were done using Bonferroni correction test. Based on Figure 5.9, the negative affect scores showed a significant decrease and increase from pre-measurement (at1) to Session 1 (at2), $p < 0.0001$. The negative affect scores then showed a significant increase and decrease from pre-measurement (t1) and Session 2 (t3), $p = 0.018$. There was no significant between Session 1 (at2) to the Session 2 (at3), $p = 1.000$.

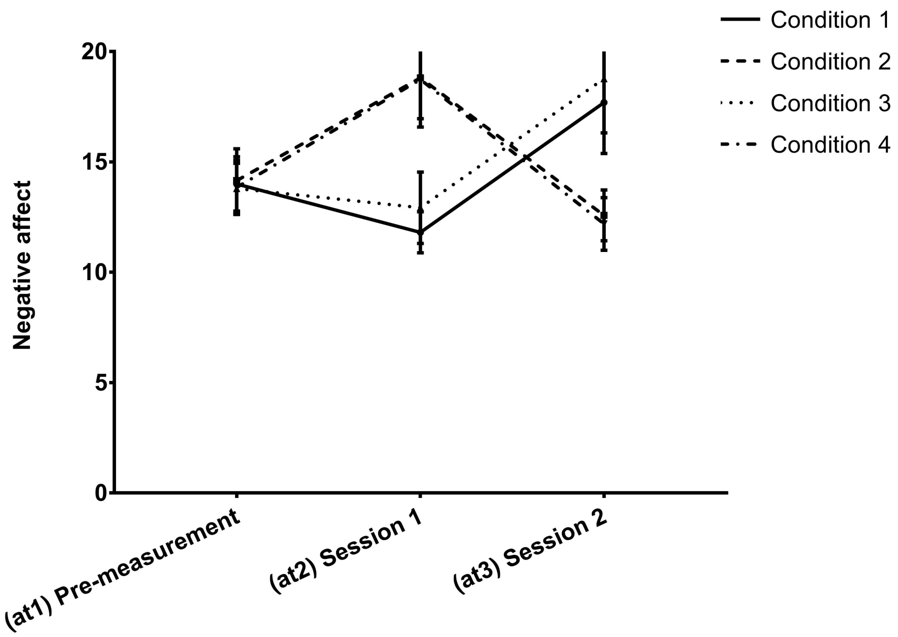


Figure 5.9: The mean with 95% confidence interval for negative affect in PANAS measures to each condition.

New variables: New variables from paired time points were created in order to check if there is difference between conditions in terms of changes in negative affect scores. The new variables are $NA\_at2mat1$ = Session 1 minus pre-measurement, $NA\_at3mat2$ = Session 2 minus Session 1 and $NA\_at3mat1$ = Session 2 minus pre-measurement. The graph for the results is shown in Figure 5.13, 216.

There was a significant difference between conditions for $NA\_at2mat1$, $F(3,98) = 26.99, p < 0.0001, r = 0.452$. The changes in negative affect scores between pre-measurement and Session 1 showed a greater decrease in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress also showed a significant decrease compared to C4:Stress_bolster, $p < 0.0001$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

There was a significant difference between conditions for $NA\_at3mat2$, $F(3,100) =$

$52.60 p < 0.0001, r = 0.612$. The changes in negative affect scores between Session 1 and Session 2 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress also showed a greater increase compared to C4:Stress_bolster, $p < 0.0001$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

There was a significant difference between conditions for $NA\_at3mat1$, $F(3,98) = 13.62, p < 0.0001, r = 0.294$. The changes in negative affect scores between pre-measurement and Session 2 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. C3:Bolster_stress also showed a greater increase compared to C4:Stress_bolster, $p < 0.0001$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

Paro or bolster session A one-way between-subjects ANOVA showed there was no significant difference between conditions at Paro and bolster session, $F(3,100) = 0.634, p = 0.595, r = 0.019$. Although no significant difference but as can be seen in Figure 5.10, the negative affect in C1:Paro_stress was lower compared to C2:Stress_Paro, C3:Bolster_stress and C4:Stress_bolster. Negative affect was higher in C3:Bolster_stress compared to other conditions.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session, $F(3,100) = 0.254, p = 0.858, r = 0.008$. Although no significant difference but as can be seen in Figure 5.10, the negative affect in C1:Paro_stress compared to other conditions.

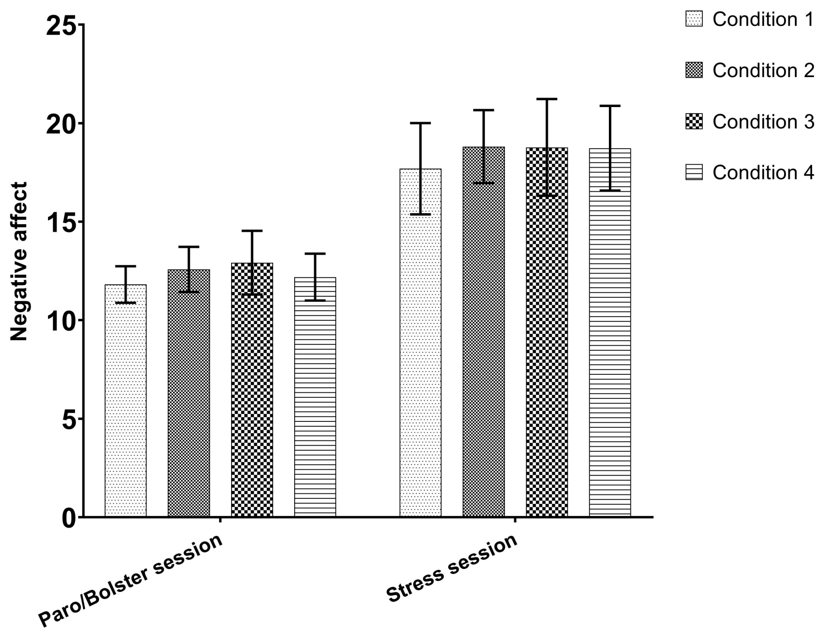


Figure 5.10: The mean with 95% confidence interval for negative affect at Paro and bolster and stress session to each condition.

Figures for psychological responses

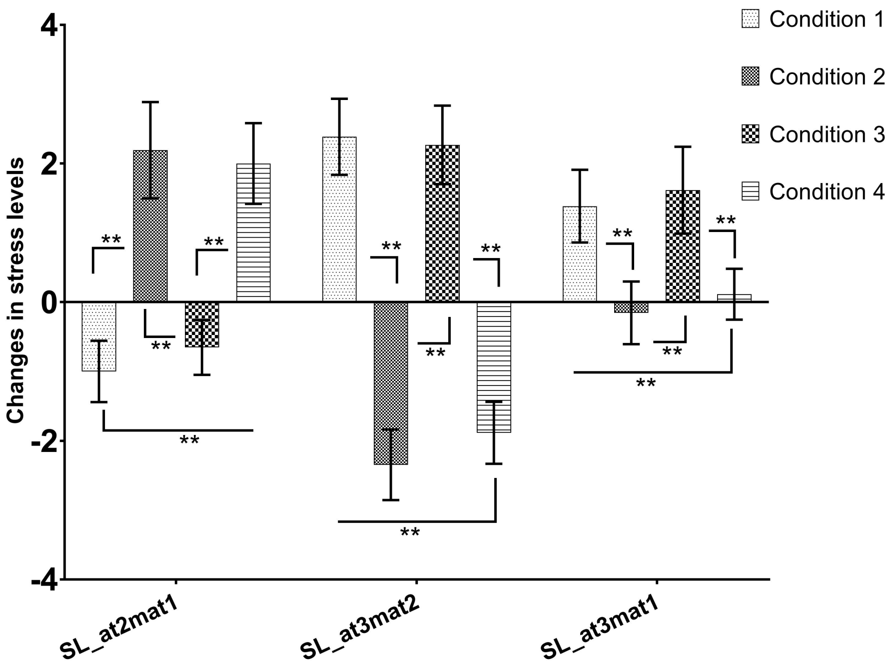


Figure 5.11: The mean with 95% confidence interval for changes of scores on levels of stress to each condition. $**\, p < 0.001$

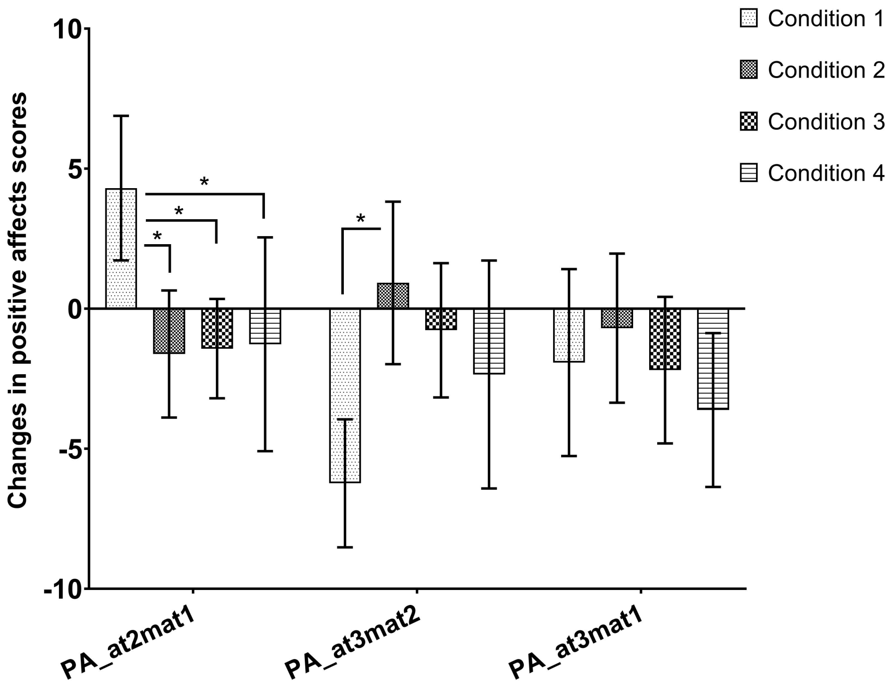


Figure 5.12: The mean with 95% confidence interval for changes in positive affect scores in PANAS for each condition. $*p < 0.05$

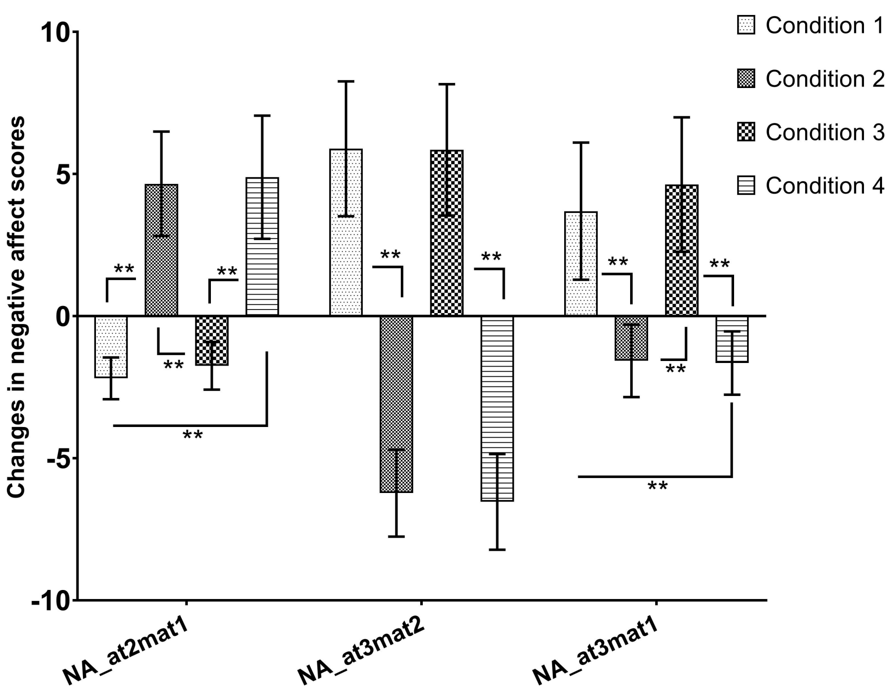


Figure 5.13: The mean with 95% confidence interval for changes in negative affect scores in PANAS for each condition. $**p < 0.001$

Self-reported emotions during Paro or bolster session

A one-way between-subjects analysis of variance showed a significant difference between conditions on self-reported of a happy emotion during Paro or bolster session, $F(3,100) = 8.608, p < 0.0001, r = 0.205$. A post hoc test using Bonferroni correction found a significant difference between C1:Paro_stress and C3:Bolster_stress, $p = 0.002$, between C1:Paro_stress and C4:Stress_bolster, $p = 0.001$, between C2:Stress_Paro and C3:Bolster_stress, $p = 0.008$ and between C2:Stress_Paro and C4:Stress_bolster, $p = 0.006$. As can be seen in Figure 5.14, the happy scores were higher in C1:Paro_stress and C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro and between C3:Bolster_stress and C4:Stress_bolster, all cases, $p = 1.000$.

A one-way between-subjects analysis of variance showed a significant difference between conditions on self-reported emotion excited during Paro or bolster session, $F(3,100) = 31.24, p < 0.0001, r = 0.484$. A post hoc test using Bonferroni correction found a significant difference between C1:Paro_stress and C3:Bolster_stress, C1:Paro_stress and C4:Stress_bolster, C2:Stress_Paro and C3:Bolster_stress and between C2:Stress_Paro and C4:Stress_bolster, all cases $p < 0.0001$. As can be seen in Figure 5.14, the excited scores were higher in C1:Paro_stress and C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 0.902$ and between C1:Paro_stress and C3:Bolster_stress, $p = 1.000$.

A one-way between-subjects analysis of variance showed no significant difference between conditions on self-reported emotion relaxed during Paro or bolster session, $F(3,100) = 2.656, p = 0.053, r = 0.074$.

A one-way between-subjects analysis of variance showed a significant difference between conditions on self-reported emotion bored during Paro or bolster session, $F(3,100) = 7.202, p < 0.0001, r = 0.178$. A post hoc test using Bonferroni correction found a significant difference between C1:Paro_stress and C3:Bolster_stress, $p = 0.002$, between C1:Paro_stress and C4:Stress_bolster, $p = 0.005$, between C2:Stress_Paro and C3:Bolster_stress, $p = 0.018$ and between C2:Stress_Paro and C4:Stress_bolster, $p = 0.039$. As can be seen in Figure 5.14, the bored scores were lower in C1:Paro_stress and C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro and between C3:Bolster_stress and C4:Stress_bolster, all cases, $p = 1.000$.

A one-way between-subjects analysis of variance showed no significant difference between conditions on self-reported emotion disgusted during Paro or bolster session, $F(3,100) = 2.437, p = 0.069, r = 0.068$. A one-way between-subjects analysis of variance showed no significant difference between conditions on self-reported emotion confused during Paro or bolster session, $F(3,100) = 1.263, p = 0.291, r = 0.036$.

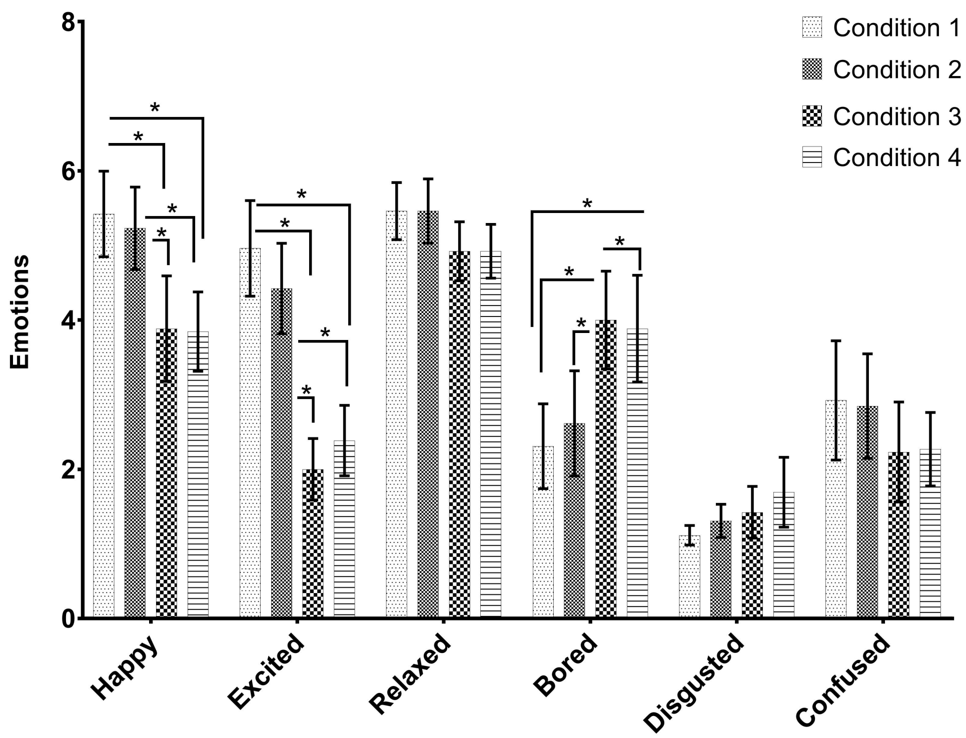


Figure 5.14: The mean with 95% confidence interval for self-reported emotions during Session 1 to each condition. $*p < 0.05$

New variables were created in order to compare the results from positive and negative emotions between conditions. The scores from happy, excited, relaxed were totalled up and the variable was named as positive emotions. The scores from bored, disgusted, confused were totalled up and the variable was named as negative emotions.

Positive expressions A one-way between-subjects analysis of variance showed a significant difference between conditions on positive emotion during Paro or bolster session, $F(3,100) = 19.52, p < 0.0001, r = 0.369$. A post hoc test using Bonferroni correction found a significant difference between C1:Paro_stress and C3:Bolster_stress, C1:Paro_stress and C4:Stress_bolster, C2:Stress_Paro and C3:Bolster_stress and between C2:Stress_Paro and C4:Stress_bolster, all cases $p < 0.0001$. As can be seen in Figure 5.15, the positive emotions scores were higher in C1:Paro_stress and C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro and between C3:Bolster_stress and C4:Stress_bolster, all cases, $p = 1.000$.

Negative expressions A one-way between-subjects analysis of variance showed no significant difference between conditions on negative emotion during Paro or bolster session, $F(3,100) = 1.310, p = 0.275, r = 0.038$.

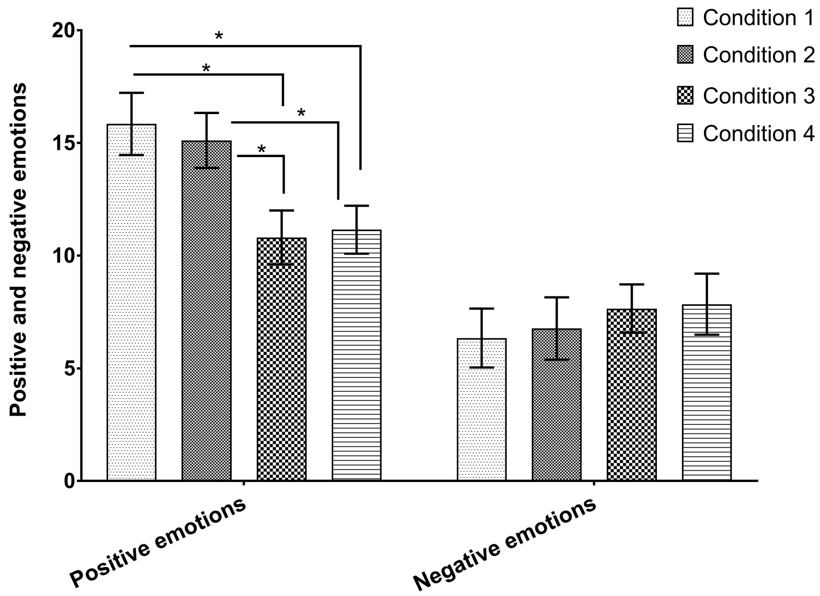


Figure 5.15: The mean with 95% confidence interval for positive and negative emotions on self-reported emotions during Session 1 to each condition. $*p < 0.05$

Self-reported noise experience

A one-way between-subjects ANOVA test showed no significant difference between conditions on noise experience (annoying), $F(3,100) = 2.212, p = 0.091, r = 0.062$. The result indicates that participants in all conditions had a similar a similar self-reported on stress experience (annoying).

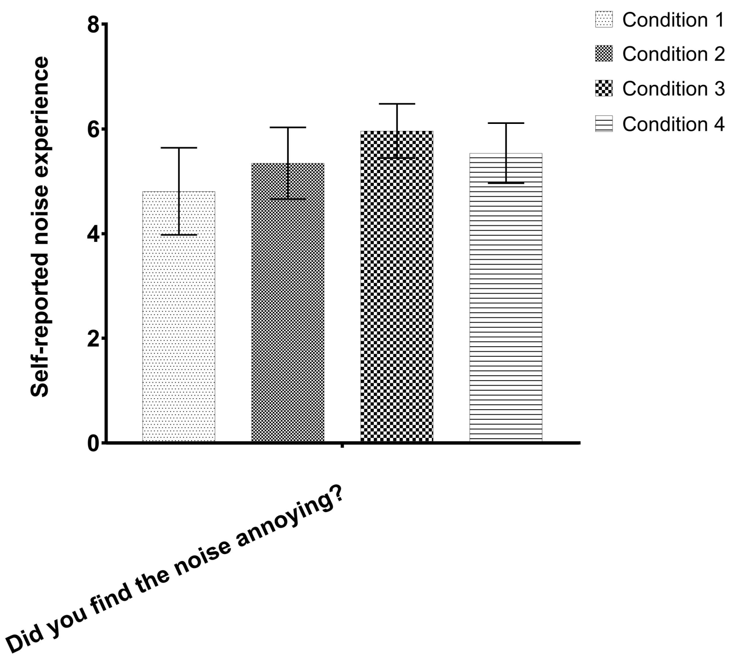


Figure 5.16: The mean with 95% confidence interval for self-reported noise experience during Session 1 to each condition.

As shown in Table 5.3, a Chi-Square test found there was no relationship between self-reported noise experience (annoying) and conditions, $p = 0.459$. Row percentages showed similar patterns for all conditions. Most participants rated the noise was extremely annoying (31.7%) to hear. At the same time, the results showed that 22.1% and 26.0% participants also rated the level of noise annoyance is above moderate.

Table 5.3: Two-way table of counts and percentages for the relationship between noise experience (annoying) and conditions

		Did you find the noise annoying?							Total
		Not at all	2	3	**Moderately**	5	6	**Extremely**	
Condition 1	Count	*3*	3	0	*2*	6	6	*6*	26
	%	11.5%	11.5%	0.0%	7.7%	23.1%	23.1%	23.1%	100.0%
Condition 2	Count	*1*	2	1	*1*	6	8	*7*	26
	%	3.8%	7.7%	3.8%	3.8%	23.1%	30.8%	26.9%	100.0%
Condition 3	Count	*0*	0	2	*2*	3	7	*12*	26
	%	0.0%	0.0%	7.7%	7.7%	11.5%	26.9%	46.2%	100.0%
Condition 4	Count	*0*	2	0	*2*	8	6	*8*	26
	%	0.0%	7.7%	0.0%	7.7%	30.8%	23.1%	30.8%	100.0%
Total	Count	*4*	7	3	*7*	23	27	*33*	104
	%	3.8%	6.7%	2.9%	6.7%	22.1%	26.0%	*31.7%*	100.0%

Perceived stress scales

A one-way between-subjects ANOVA showed no significant difference between conditions on perceived stress scores at pre-measurement (at1), $F(3,100) = 0.291, p = 0.832, r = 0.009$. The result indicates that participants in all conditions had a similar level of perceived stress at the beginning of study.

Then a comparison between pre-measurement and post-measurement of perceived stress was done using a mixed between-within subjects ANOVA. The result showed no significant effect of time on perceived stress, $F(1,100) = 1.045, p = 0.309, r = 0.010$ and no significant effect of conditions, $F(3,100) = 0.626, p = 0.600, r = 0.018$. There was no significant interaction between condition and time on perceived stress, $F(3,100) = 1.351, p = 0.262, r = 0.039$. There was no linear trend to perceived stress scores, $F(1,100) = 1.045, p = 0.309, r = 0.010$.

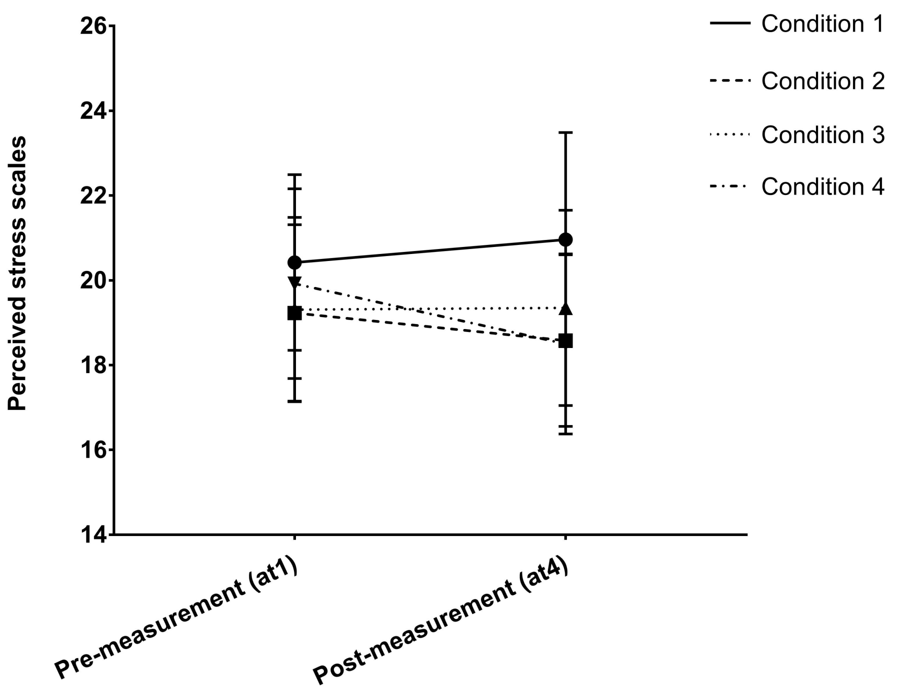


Figure 5.17: The mean with 95% confidence interval for perceived stress scales to each condition.

Further analysis was done to determine relationship between PSS-10 pre-measurement and post-measurement (at4) to each condition. The PSS-10 scores for each time-point were split into two categories, *Low stress* and *High stress*. Participants with total scores that were higher than 19 were considered as high level of stress (Leroy & Miller 2010, Lovell et al. 2011).

McNemar test found there was a significant difference in PSS-10 scores between pre-measurement and post-measurement (at4) in C1:Paro_stress, $p = 0.016$. As shown in Table 5.4, based on total percentages in C1:Paro_stress, the percentages of low stress increase from 57.7.0% at pre-measurement to 84.6% at post-measurement (at4). However, the percentages of high stress reduce from 42.3% at pre-measurement to 15.4% post-measurement (at4). There were no significant changes in PSS-10 scores between pre-measurement and post-measurement (at4) in C2:Stress_Paro, C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$. Although the results were not significant, for C2:Stress_Paro, the percentages for low stress scores were reduce from 65.4% at pre-measurement to 61.5% post-measurement (at4). However, for C4:Stress_bolster, the percentages for low stress scores were increase from 57.7% at pre-measurement to 61.5% post-measurement (at4). There was no change in PSS-10 scores for C3:Bolster_stress at pre-measurement and post-measurement (at4).

Table 5.4: Two-way table of counts and percentages for the relationship between PSS-10 pre- and post-measurement

PSS-10 groups		Condition			
		Condition 1	Condition 2	Condition 3	Condition 4
Low	Count	15	17	12	15
	Pre	57.7%	65.4%	46.2%	57.7%
	Count	22	16	12	16
	Post	84.6%	61.5%	46.2%	61.5%
High	Count	11	9	14	11
	Pre	42.3%	34.6%	53.8%	42.3%
	Count	4	10	14	10
	Post	15.4%	38.5%	53.8%	38.5%

5.9.3 Effects of conditions on observed facial emotional expressions

The 5 minutes of video recording during the Paro and bolster sessions were examined for the number of occurrences of four facial expressions: positive expressions (smile and laugh) and negative expressions (boredom and disgust). The 5 minutes of videos were divided into 30 blocks of 10 seconds.

All the videos were analysed by two evaluators for inter rater reliability. The inter rater reliability for all facial expressions were analysed using Intraclass Correlation Coefficient (ICC). A moderate degree of reliability was found between two evaluators for all facial expressions, α between 0.733 to 0.894.

Then after checking the inter rater reliability, the number of coded facial expression scores were totalled for each emotion for each condition, using the mean scores of two evaluators for further analysis.

A one-way between-subjects ANOVA test was used to test for differences between conditions for each coded expression: smile, laugh, disgust and boredom. There was a significant difference between conditions in the number of times each of the following expressions occurred: smile, $F(3,100) = 77.79, p < 0.0001, r = 0.700$, laugh, $F(3,100) = 11.28, p < 0.0001, r = 0.253$ and boredom, $F(3,100) = 172.77, p < 0.0001, r = 0.838$. There was no significant difference between condition in C3:Bolster_stress, for coded facial expression, disgust, $F(3,100) = 0.968, p = 0.411, r = 0.028$. The graph for the results is shown in Figure 5.18.

For the coded facial expression, smile, a post-hoc comparison found a significant difference between C1:Paro_stress and C3:Bolster_stress, $p < 0.0001$, between C1:Paro_stress and C4:Stress_bolster, $p < 0.0001$, between C2:Stress_Paro and C3:Bolster_stress, $p < 0.0001$ and between C2:Stress_Paro and C4:Stress_bolster, $p < 0.0001$. The smiling scores were higher in C1:Paro_stress compared to C3:Bolster_stress and C4:Stress_bolster. The smiling scores also higher in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster.

There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 1.000$ and C3:Bolster_stress and C4:Stress_bolster, $p = 1.000$.

For the coded facial expression, laugh, a post-hoc comparison found significant difference between C1:Paro_stress and C3:Bolster_stress, $p = 0.003$, between C1:Paro_stress and C4:Stress_bolster, $p = 0.002$, between C2:Stress_Paro and C3:Bolster_stress, $p < 0.0001$ and between C2:Stress_Paro and C4:Stress_bolster, $p < 0.0001$. The laughing scores were higher in C1:Paro_stress compared to C3:Bolster_stress and C4:Stress_bolster. The laughing scores also higher in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 1.000$ and C3:Bolster_stress and C4:Stress_bolster, $p = 1.000$.

For the coded facial expression, boredom, post-hoc comparison found a significant difference between C1:Paro_stress and C3:Bolster_stress, $p < 0.0001$, between C1:Paro_stress and C4:Stress_bolster, $p < 0.0001$, between C2:Stress_Paro and C3:Bolster_stress, $p < 0.0001$ and between C2:Stress_Paro and C4:Stress_bolster, $p < 0.0001$. The boredom scores were lower in C1:Paro_stress compared to C3:Bolster_stress and C4:Stress_bolster. The boredom scores also lower in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 1.000$ and C3:Bolster_stress and C4:Stress_bolster, $p = 0.147$.

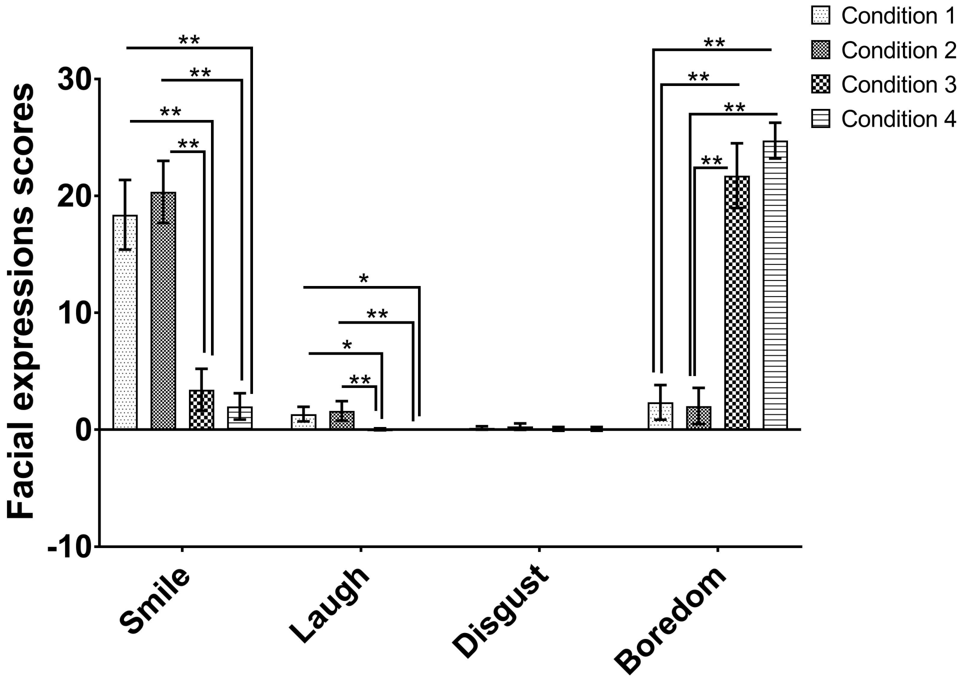


Figure 5.18: The mean with 95% confidence interval for facial emotional expressions to each condition. $*p < 0.05, **p < 0.001$

A further analysis was done to compare scores for positive and negative facial expressions for each condition. A new variable, positive facial expressions, was created by combining the smile and laugh expression scores. Another variable, negative facial expression, was formed by combining the disgust and boredom expression scores.

Positive expressions A one-way between-subjects ANOVA test found a significant difference between conditions for positive facial expressions, $F(3,100) = 79.54, p < 0.0001, r = 0.705$. A post hoc test using Bonferroni correction was carried out. As shown in Figure 5.19, there was a significant difference between C1:Paro_stress and C3:Bolster_stress, $p < 0.0001$, between C1:Paro_stress and C4:Stress_bolster, $p < 0.0001$, between C2:Stress_Paro and C3:Bolster_stress, $p < 0.0001$ and, between C2:Stress_Paro and C4:Stress_bolster, $p < 0.0001$. The positive expressions scores were higher in C1:Paro_stress compared to C3:Bolster_stress and C4:Stress_bolster. It was found that the positive expressions scores were also higher in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 1.000$, and between C3:Bolster_stress and C4:Stress_bolster, $p = 1.000$.

Negative expressions A one-way between-subjects ANOVA test found a significant difference between conditions for negative facial expressions, $F(3,100) = 167.22, p < 0.0001, r = 0.834$. A post hoc test using Bonferroni correction was carried out. As shown in Figure 5.19, there was a significant difference between C1:Paro_stress and C3:Bolster_stress, $p < 0.0001$, between C1:Paro_stress and C4:Stress_bolster, $p < 0.0001$, between C2:Stress_Paro and C3:Bolster_stress, $p < 0.0001$ and, between C2:Stress_Paro and C4:Stress_bolster, $p < 0.0001$. The negative expressions scores were lower in C1:Paro_stress compared to C3:Bolster_stress and C4:Stress_bolster. It was found that the negative expressions scores were also lower in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, $p = 1.000$ and between C3:Bolster_stress and C4:Stress_bolster, , $p = 0.155$.

Paired t-test found there was a significant difference in the scores between positive and negative facial expressions at each condition, in C1:Paro_stress, $t(25) = 8.472, p < 0.0001$, in C2:Stress_Paro, $t(25) = 10.21, p < 0.0001$, in C3:Bolster_stress, $t(25) = -8.546, p < 0.0001$ and C4:Stress_bolster, $t(25) = -20.23, p < 0.0001$. The results showed positive expressions scores were higher in C1:Paro_stress and C2:Stress_Paro compared to negative facial expressions. However, the negative expression scores were higher in C3:Bolster_stress and C4:Stress_bolster compared to positive expression scores.

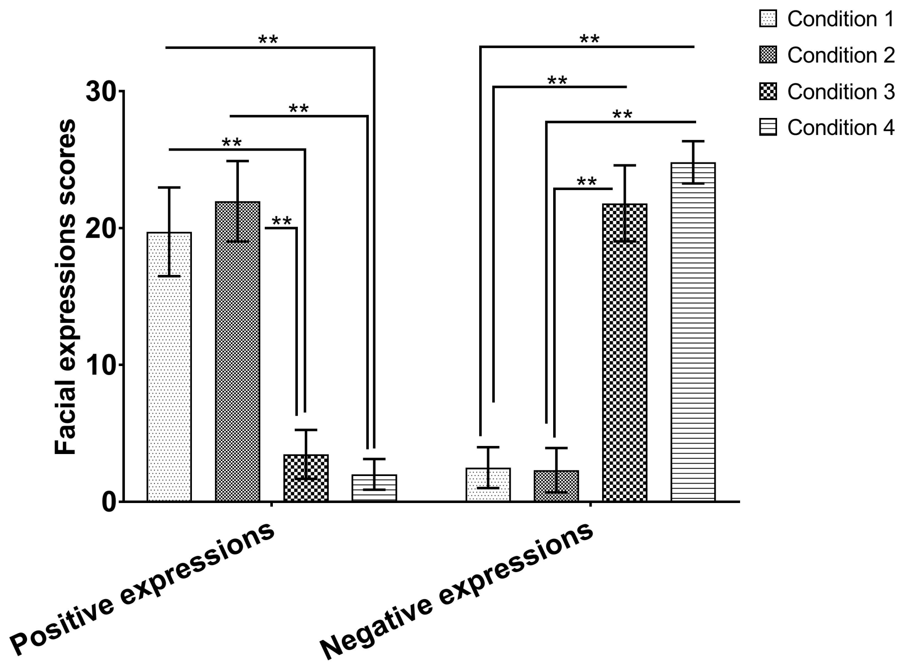


Figure 5.19: The mean with 95% confidence interval for negative and positive expressions to each condition. $**p < 0.001$

5.9.4 Mathematical task scores

A one-way between-subjects ANOVA showed no significant difference between conditions on mathematical task scores, $F(3,100) = 0.147, p = 0.932, r = 0.004$.

Further analysis using Chi-square to analyse the relationship between categories of mathematical task scores (low, high) and conditions. The mathematical task scores for each timepoint were split into two categories, *Low* and *High*. Participants with total scores that higher than 5 were considered as high. There was no significant difference between categories of mathematical task scores at pre-measurement (low, high) and conditions, $X^2(3, N = 104) = 0.707, p = 0.857$.

Table 5.5: Two way table of counts and percentages for the relationship between categories of mathematical scores (low, high) and conditions

Mathematical scores categories		Condition 1	Condition 2	Condition 3	Condition 4
Low	Count	13	12	14	11
	% within Condition	50.0%	46.2%	53.8%	42.3%
High	Count	13	14	12	15
	% within Condition	50.0%	53.8%	46.2%	57.7%

But as can be seen in summaries of cross-tabulation Table 5.5, the number of participants in low and high scores categories were equal in C1:Paro_stress, but high group were lower in bolster condition. However, C2:Stress_Paro and C4:Stress_bolster showed a higher number of participants in high scores category.

5.9.5 Effects of conditions on physiological responses

Skin conductance responses

A mixed between-within subjects analysis of variance ANOVA showed there was a significant effect of time on SCR, $F(3,300) = 41.98, p < 0.0001, r = 0.296$, and no significant effect of conditions, $F(3,100) = 1.683, p = 0.175, r = 0.048$. There was significant interaction between condition and time on SCR, $F(9,300) = 2.772, p = 0.004, r = 0.077$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. As evident in Figure 5.20, there was a significant increase in SCR from pre-measurement (t1) to Session 1 (t2), $p < 0.0001$. There was a significant decrease in SCR from Session 2 (t3) to post-measurement (t4), $p < 0.0001$. However, there was no significant change from pre-measurement (t1) to post-measurement (t4), $p = 0.849$ and Session 1 (t2) to Session 2 (t3), $p = 0.389$.

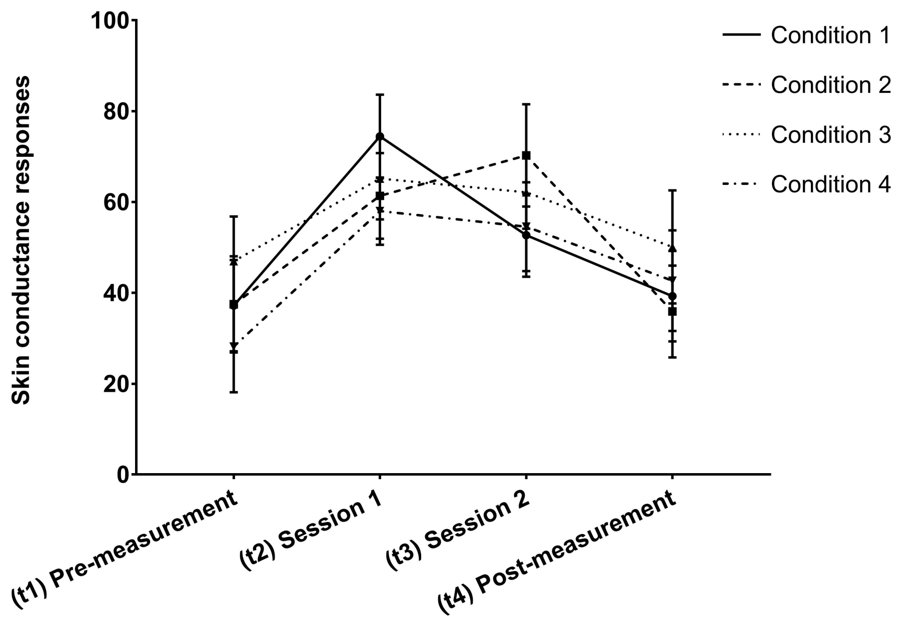


Figure 5.20: The mean with 95% confidence interval for skin conductance responses to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in SCR. The new variables are SCR_t2mt1 = Session 1 minus pre-measurement, SCR_t3mt2 = Session 2 minus Session 1 and SCR_t4mt3 = post-measurement minus Session 2, and SCR_t4mt1 = post-measurement minus pre-measurement.

There was no significant difference between conditions for SCR_t2mt1, $F(3,100) = 2.202, p = 0.093, r = 0.063$. As is apparent in Figure 5.28 on page 242, all condition showed an increase in SCR.

There was a significant difference between conditions for SCR_t3mt2, $F(3,100) = 6.029, p = 0.001, r = 0.153$. The changes in SCR scores between Session 1 and Session 2 showed a greater decrease in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$. Furthermore, as can be seen in Figure 5.28 on page 242, all conditions showed a reduction in SCR. C3:Bolster_stress and C4:Stress_bolster showed a low reduction in SCR.

There was a significant difference between conditions for and SCR_t4mt3, $F(3,100) = 3.931, p = 0.011, r = 0.105$. The changes in SCR scores between Session 2 and post-measurement showed a significant difference between C2:Stress_Paro and C4:Stress_bolster, $p = 0.030$. As can be seen in Figure 5.28 on page 242, other conditions showed an increase in SCR, and C4:Stress_bolster showed a greater increase compared to other conditions. Only C2:Stress_Paro showed a small reduction in SCR.

There was no significant difference between conditions for SCR_t4mt1, $F(3,100) = 1.287, p = 0.283, r = 0.037$. Although no significant difference in SCR changes scores for SCR_t4mt1, as can be seen in Figure 5.28 on page 242, there was a greater reduction in SCR in C1:Paro_stress compared to other conditions. There was also slightly increase in SCR in C2:Stress_Paro compared to C3:Bolster_stress and C4:Stress_bolster.

Paro or bolster session A one-way between-subjects ANOVA showed there was a significant difference between conditions at Paro and bolster session, $F(3,100) = 3.231, p = 0.026, r = 0.088$. But, only SCR between C1:Paro_stress and C4:Stress_bolster was significant differ, $p = 0.024$. However, other planned comparison condition showed no significant difference. Although no significant difference but as can be seen in Figure 5.21, C1:Paro_stress and C2:Stress_Paro which interacting with Paro showed higher in SCR compared to bolster sessions C3:Bolster_stress and C4:Stress_bolster.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session, $F(3,100) = 1.075, p = 0.363, r = 0.031$. Although no significant difference but as can be seen in Figure 5.21, the SCR in C1:Paro_stress was lower compared to other conditions, especially when comparing it with bolster session (C3:Bolster_stress).

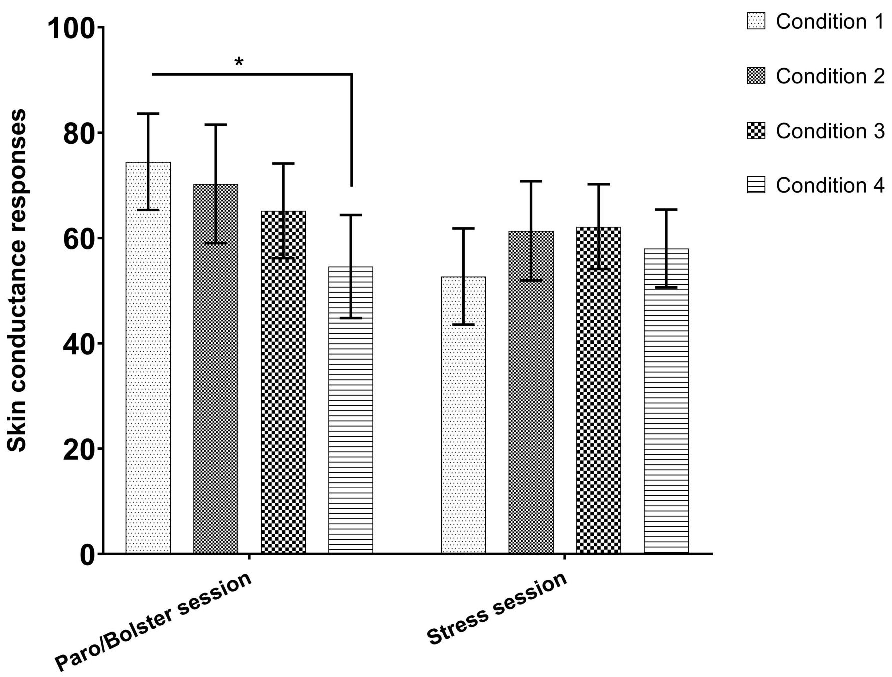


Figure 5.21: The mean with 95% confidence interval for skin conductance at Paro and bolster and stress session to each condition. $*p < 0.05$.

Heart rate

A mixed between-within subjects analysis of variance ANOVA showed there was a significant effect of time on HR, $F(2.409, 240.86) = 30.56, p < 0.0001, r = 0.234$, and no significant effect of conditions, $F(3, 100) = 0.061, p = 0.980, r = 0.002$. There was significant interaction between condition and time on HR, $F(7.226, 240.86) = 10.03, p < 0.0001, r = 0.231$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. There was a significant increase in HR from pre-measurement (t1) to Session 1 (t2), $p < 0.0001$. There was a significant decrease in HR from Session 2 (t3) to post-measurement (t4), $p < 0.0001$. (Refer to Figure 5.22). However, there was no significant change from pre-measurement (t1) to post-measurement (t4), $p = 0.245$ and Session 1 (t2) to Session 2 (t3), $p = 0.133$.

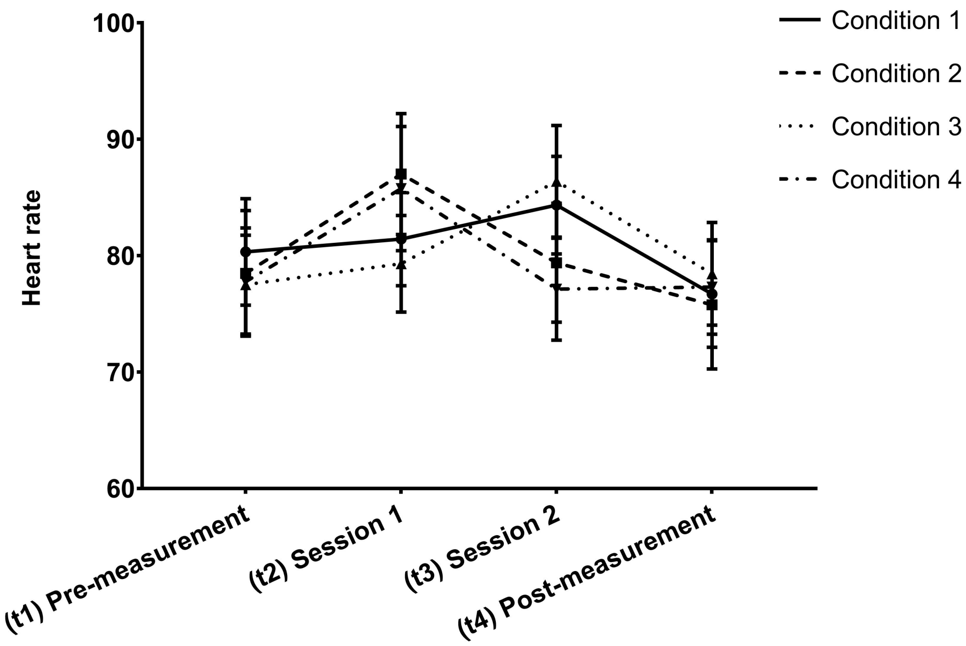


Figure 5.22: The mean with 95% confidence interval for heart rate to each condition. $*p < 0.05$

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in HR. The new variables are HR_t2mt1 = Session 1 minus pre-measurement, HR_t3mt2 = Session 2 minus Session 1 and HR_t4mt3 = post-measurement minus Session 2, and HR_t4mt1 = post-measurement minus pre-measurement.

There was a significant difference between conditions for HR_t2mt1, $F(3,100) = 2.202, p = 0.093, r = 0.063$. As is apparent in Figure 5.29 on page 243, all condition showed an increase in HR. But, C2:Stress_Paro showed a greater increase in changes score of HR compared to C1:Paro_stress, $p = 0.002$. C4:Stress_bolster also showed a greater increase in changes score of HR compared to C3:Bolster_stress, $p = 0.016$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$.

There was a significant difference between conditions for HR_t3mt2, $F(3,100) =$

$33.35, p < 0.0001, r = 0.500$. As evident in Figure 5.29 on page 243, the changes in HR scores between Session 1 and Session 2 showed a greater increase in C1:Paro_stress compared to C2:Stress_Paro, $p < 0.0001$ and, between C3:Bolster_stress and C4:Stress_bolster, $p < 0.0001$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

There was a significant difference between conditions for and HR_t4mt3, $F(3,100) = 5.770, p = 0.001, r = 0.148$. As can be seen in Figure 5.29 on page 243, the changes in HR scores between Session 2 and post-measurement showed a greater decrease in C3:Bolster_stress compared to C4:Stress_bolster, $p = 0.003$. There was no significant difference between C1:Paro_stress and C2:Stress_Paro, between C1:Paro_stress and C3:Bolster_stress and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

There was no significant difference between conditions for HR_t4mt1, $F(3,100) = 2.172, p = 0.096, r = 0.061$. Although no significant difference in HR changes scores for HR_t4mt1, as can be seen in Figure 5.29 on page 243, there was a little reduction in HR in C1:Paro_stress, C2:Stress_Paro and C4:Stress_bolster. There was also slightly increase in HR in C3:Bolster_stress.

Paro or bolster session A one-way between-subjects ANOVA showed there was no significant difference between conditions at Paro or bolster session, $F(3,100) = 0.666, p = 0.575, r = 0.020$.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session, $F(3,100) = 0.231, p = 0.875, r = 0.007$.

Taken together both session, although no significant difference for both time points, but as can be seen in Figure 5.23, the HR in C1:Paro_stress was lower compared to other conditions during stress induction. Although, the HR was slightly higher in C1:Paro_stress during the Paro or bolster session.

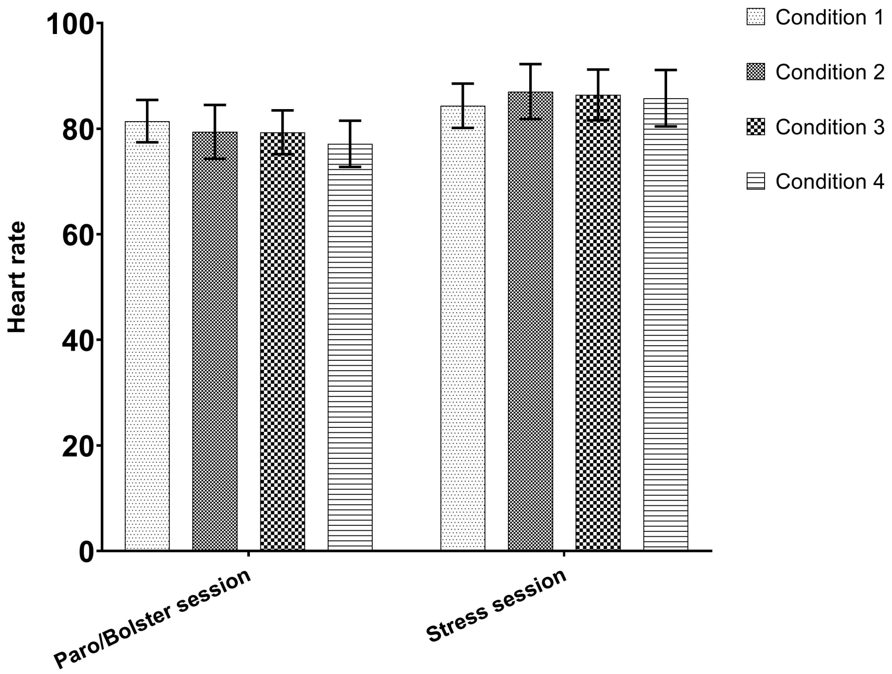


Figure 5.23: The mean with 95% confidence interval for heart rate at Paro and bolster and stress session to each condition.

Standard of all normal RR intervals

7 outliers were removed from data analysis.

A mixed between-within subjects ANOVA showed there was a significant effect of time on SDNN, $F(3,291) = 7.929, p < 0.0001, r = 0.076$, and no significant effect of conditions, $F(3,97) = 0.066, p = 0.978, r = 0.002$. There was significant interaction between condition and time on SDNN, $F(9,291) = 2.053, p = 0.034, r = 0.060$.

Planned pairwise comparisons between time points were done using Bonferroni correction test. As can be seen in Figure 5.24, there was a significant decrease in SDNN from pre-measurement (t1) to Session 1 (t2), $p = 0.022$. There was also a significant decrease in SDNN from Session 2 (t3) to post-measurement (t4), $p = 0.039$. However, there was no significant change from pre-measurement (t1) to post-measurement (t4), $p = 0.659$ and Session 1 (t2) to Session 2 (t3), $p = 0.121$.

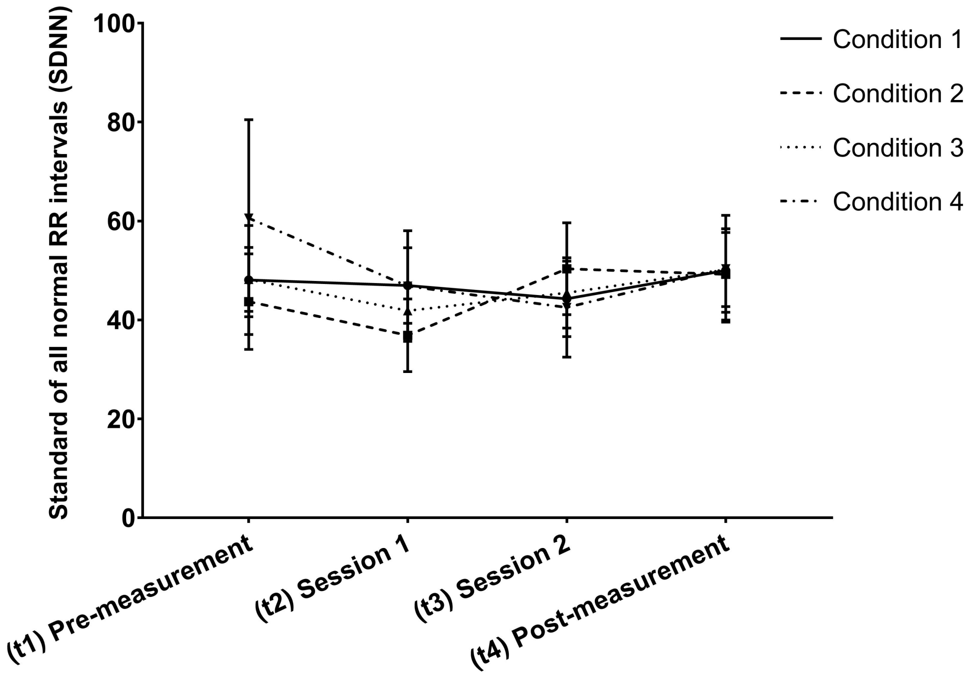


Figure 5.24: The mean with 95% confidence interval for SDNN to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in SDNN. The new variables are $SDNN\_t2mt1$ = Session 1 minus pre-measurement, $SDNN\_t3mt2$ = Session 2 minus Session 1 and $SDNN\_t4mt3$ = post-measurement minus Session 2, and $SDNN\_t4mt1$ = post-measurement minus pre-measurement.

There was no significant difference between conditions for $SDNN\_t2mt1$, $F(3,98) = 0.836, p = 0.477, r = 0.025$. As is apparent in Figure 5.30 on page 244, all conditions showed a decrease in SDNN. But, C1:Paro_stress showed a small decrease in SDNN compared to other conditions.

There was a significant difference between conditions for $SDNN\_t3mt2$, $F(3,97) = 6.405, p = 0.001, r = 0.165$. As evident in Figure 5.30 on page 244, the changes in SDNN scores between Session 1 and Session 2 showed a greater decrease in C1:Paro_stress compared to C2:Stress_Paro, $p = 0.001$. There was also a greater increase in SDNN in C2:Stress_Paro compared to C4:Stress_bolster, $p = 0.005$. There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$.

There was no significant difference between conditions for and $SDNN\_t4mt3$, $SDNN\_t4mt3$, $F(3,98) = 1.623, p = 0.189, r = 0.047$. Although no significant difference in SDNN changes scores, but, as can be seen in Figure 5.30 on page 244, other conditions showed an increase in SDNN, and C2:Stress_Paro showed a small decrease in SDNN.

There was no significant difference between conditions for $SDNN\_t4mt1$, $F(3,97) = 0.384, p = 0.764, r = 0.012$. Although no significant difference in SDNN changes scores for $SDNN\_t4mt1$, as can be seen in Figure 5.30 on page 244, all conditions showed an increase in SDNN. There was also a greater increase in SDNN in C2:Stress_Paro compared to other conditions.

Paro or bolster session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session for SDNN, $F(3,98) = 1.014, p = 0.390, r = 0.030$. Although no significant difference but as can be seen in Figure 5.25, the SDNN in C3:Bolster_stress and C4:Stress_bolster were lower compared to C1:Paro_stress and C2:Stress_Paro.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session for SDNN, $F(3,98) = 1.208, p = 0.311, r = 0.036$. Although no significant difference but as can be seen in Figure 5.25, the SDNN in C2:Stress_Paro was lower compared to other conditions.

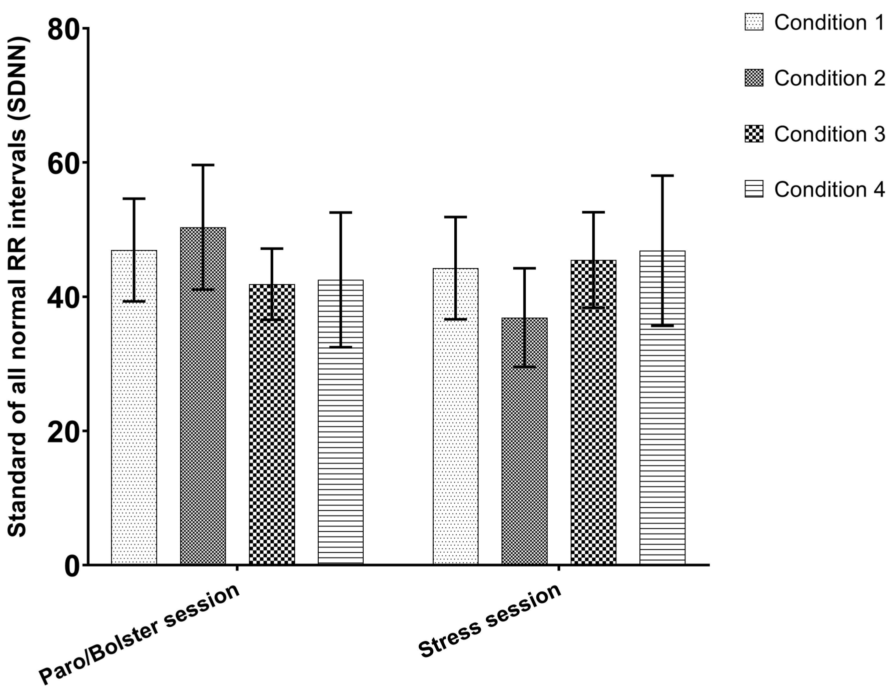


Figure 5.25: The mean with 95% confidence interval for SDNN at Paro and bolster and stress session to each condition.

High frequency

37 outliers were removed from data analysis.

A mixed between-within subjects analysis of variance ANOVA showed there was no significant effect of time on HF, $F(2.125, 165.73) = 2.297, p = 0.100, r = 0.029$, but a significant effect of conditions, $F(3,78) = 6.635, p < 0.0001, r = 0.203$. There was a significant interaction between condition and time on HF, $F(6.374, 165.73) = 4.783, p < 0.0001, r = 0.155$.

A time trend analysis showed a significant linear time trend, $F(1,78) = 4.654, p = 0.034, r = 0.056$.

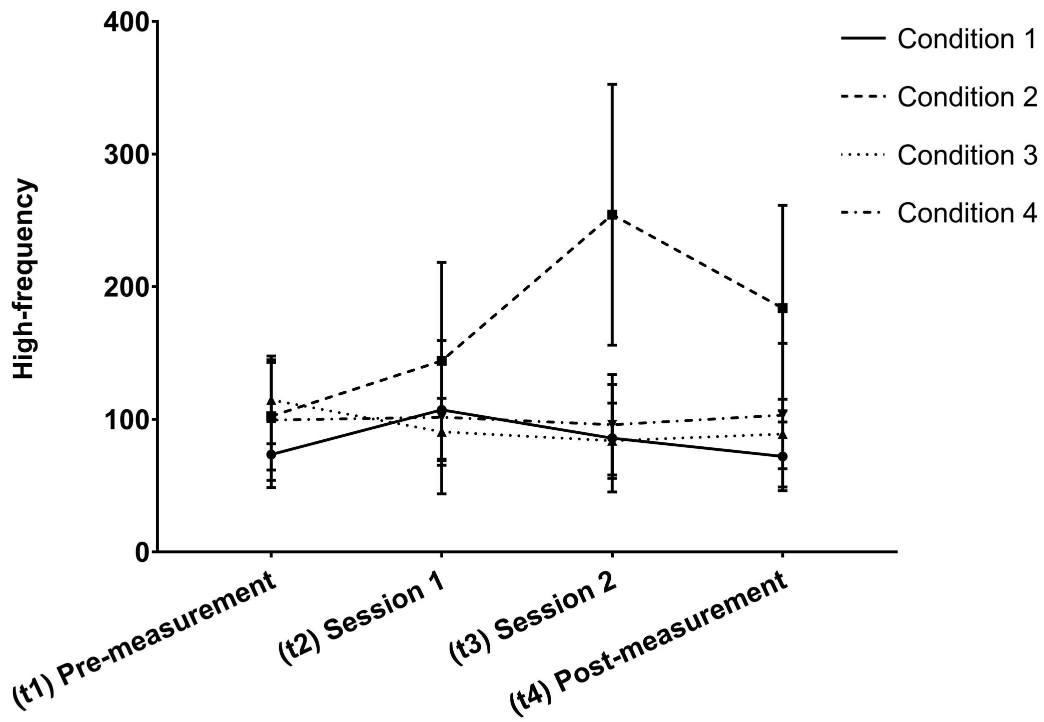


Figure 5.26: The mean with 95% confidence interval for high-frequency to each condition.

New variables: New variables were created in order to check if there were differences between conditions in terms of changes in HF. The new variables are HF_t2mt1 = Session 1 minus pre-measurement, HF_t3mt2 = Session 2 minus Session 1 and HF_t4mt3 = post-measurement minus Session 2, and HF_t4mt1 = post-measurement minus pre-measurement.

There was no significant difference between conditions for HF_t2mt1, $F(3,86) = 0.937, p = 0.427, r = 0.032$. Although not significant, as is apparent in Figure 5.31 on page 245, C1:Paro_stress and C2:Stress_Paro showed an increase in HF, and C3:Bolster_stress and C4:Stress_bolster showed a decrease.

There was a significant difference between conditions for HF_t3mt2, $F(3,85) = 6.126, p = 0.001, r = 0.178$. The changes in HF scores between Session 1 and Session 2 showed a greater increase in C2:Stress_Paro compared to C1:Paro_stress, $p = 0.009$, and C4:Stress_bolster,

$p = 0.021$. (Refer to Figure 5.31 on page 245). There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$.

There was no significant difference between conditions for and HF_t4mt3, $F(3,85) = 2.392, p = 0.074, r = 0.078$. Although no significant difference in HF changes scores for HF_t4mt3, as can be seen in Figure 5.31 on page 245, there was a greater reduction in HF in C2:Stress_Paro compared to other conditions. Only C4:Stress_bolster showed a small increase in HF.

There was a significant difference between conditions for HF_t4mt1, $F(3,86) = 6.619, p < 0.0001, r = 0.188$. The changes in HF scores between pre-measurement and post-measurement showed a greater increase in C2:Stress_Paro compared to C1:Paro_stress, $p = 0.012$. (Refer to Figure 5.31 on page 245). There was no significant difference between C1:Paro_stress and C3:Bolster_stress, between C3:Bolster_stress and C4:Stress_bolster, and between C2:Stress_Paro and C4:Stress_bolster, all cases $p = 1.000$.

Paro or bolster session A one-way between-subjects ANOVA showed there was a significant difference between conditions at Paro or bolster session, $F(3,93) = 7.814, p < 0.0001, r = 0.201$. The HF in C2:Stress_Paro was higher compared to C1:Paro_stress, $p = 0.002$ and C4:Stress_bolster, $p = 0.001$. (Refer to Figure 5.27). There was no significant difference between C1:Paro_stress and C3:Bolster_stress, and between C3:Bolster_stress and C4:Stress_bolster, all cases $p = 1.000$.

Stress session A one-way between-subjects ANOVA showed there was no significant difference between conditions at stress session, $F(3,94) = 1.227, p = 0.304, r = 0.038$. Although no significant difference but as can be seen in Figure 5.27, the HF in C2:Stress_Paro was higher compared to C1:Paro_stress. The HF in C4:Stress_bolster was higher compared to C3:Bolster_stress.

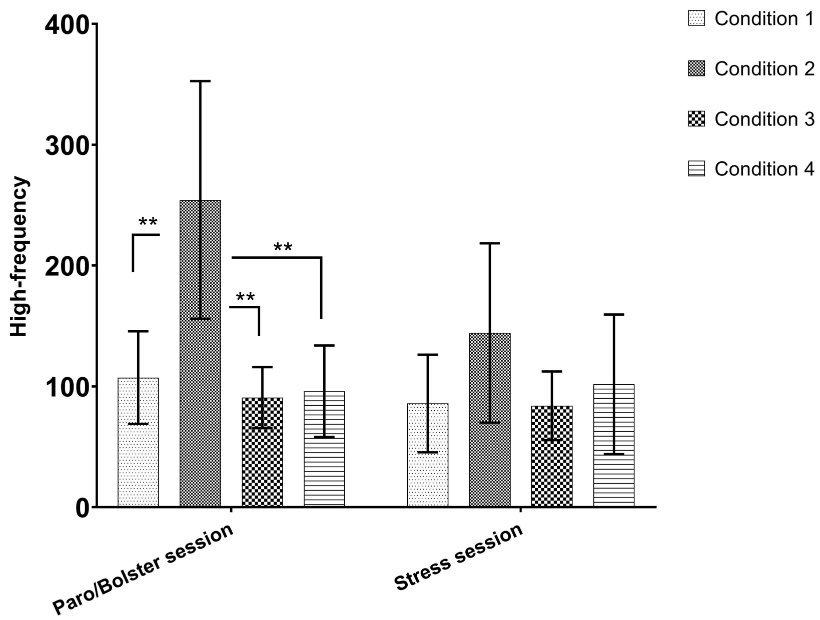


Figure 5.27: The mean with 95% confidence interval for high-frequency at Paro and bolster and stress session to each condition. $**p < 0.001$

Figures changes in scores for physiological responses

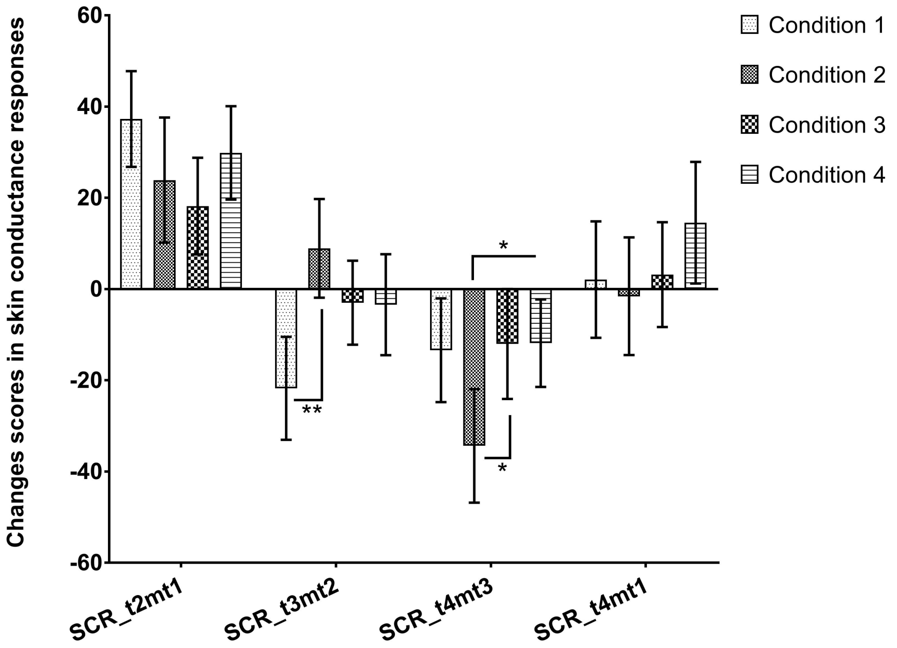


Figure 5.28: The mean with 95% confidence interval for changes in scores of skin conductance responses to each condition. $*p < 0.05, **p < 0.001$

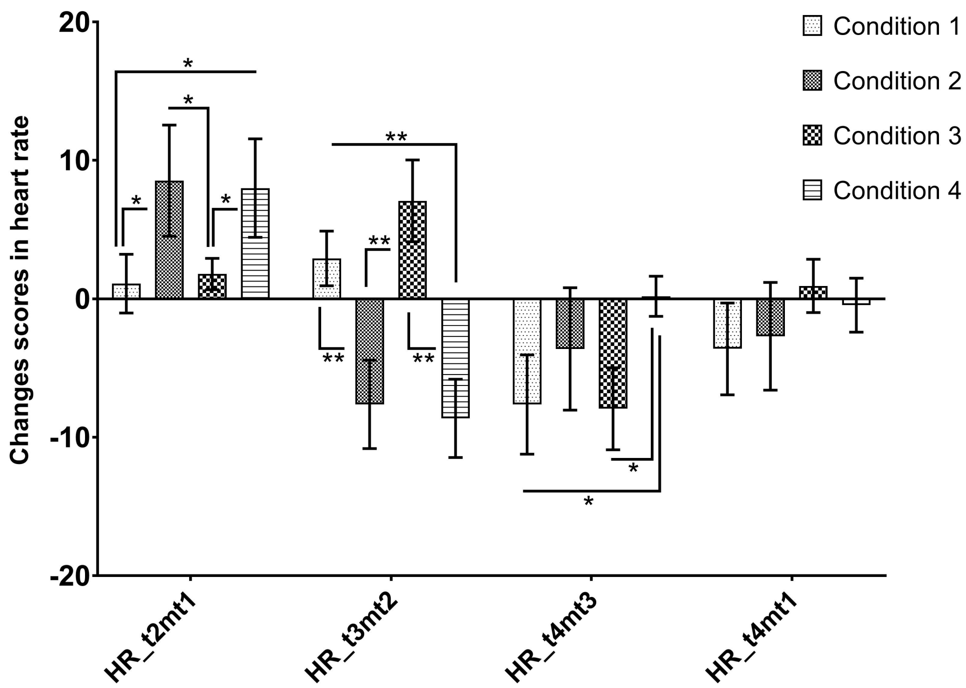


Figure 5.29: The mean with 95% confidence interval for changes in scores of heart rate to each condition. $*p < 0.05, **p < 0.001$

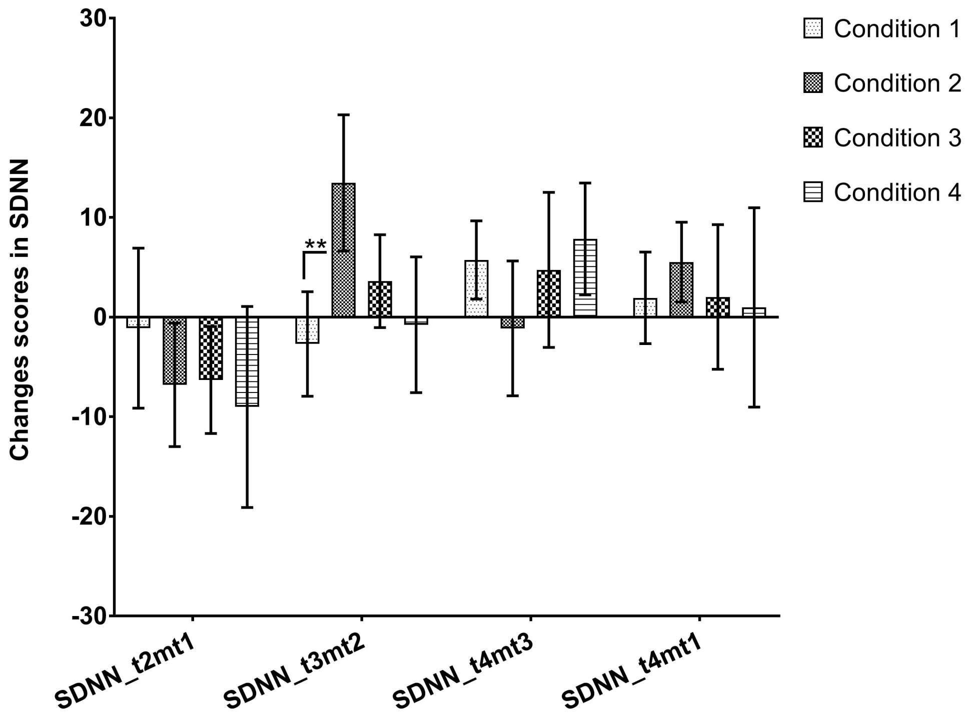


Figure 5.30: The mean with 95% confidence interval for changes in scores of SDNN to each condition. $**p < 0.001$

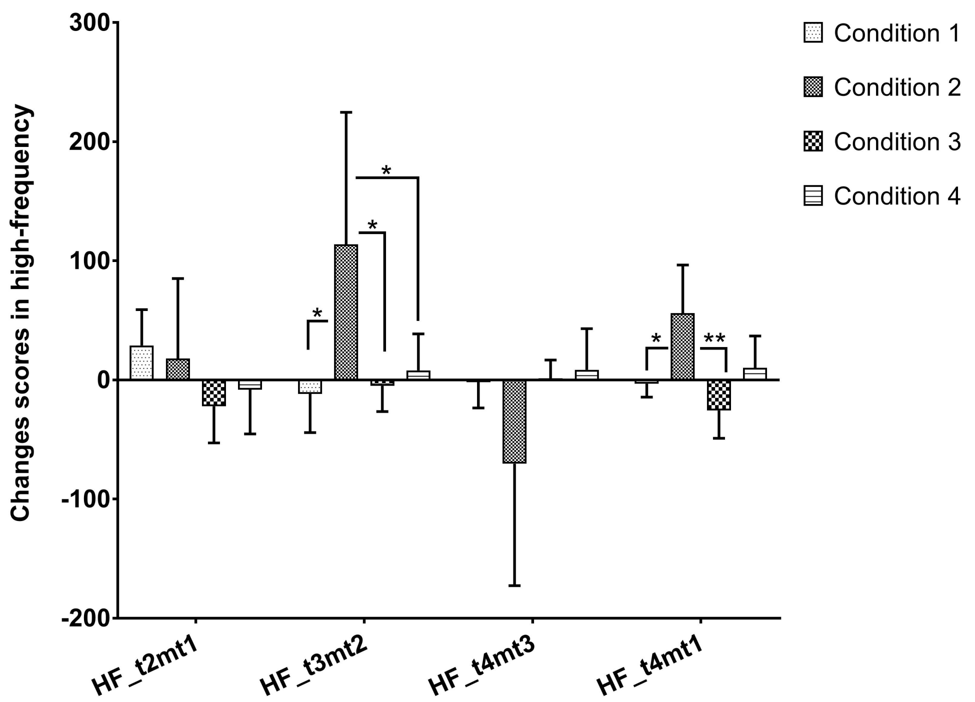


Figure 5.31: The mean with 95% confidence interval for changes in scores of high-frequency to each condition. $*p < 0.05, **p < 0.001$

5.10 Discussion

The general aim of Study 3 was to investigate the effect of short-term interaction with Paro on psychological and physiological responses, and the potential of the interaction for reducing the effects of induced stress. The short-term intervention and its effect on induced stress was also compared in different intervention formats. At the same time, Study 3 examined and compared the immediate effects of a Paro session and a bolster session. The effects were measured using self-reported stress, emotions, negative and positive affect, perceived stress, and physiological responses: skin conductance response (SCR), heart rate (HR), and heart rate variability (HRV), and observed facial observation, and scores in a mathematical task. Additional results, self-reported of math anxiety and noise experience.

5.10.1 Effects of conditions on psychological responses at the end of experiment compared to baseline

This section focuses on hypotheses H1(i)(a) and H1(ii)(a) and based on the results from psychological responses based on changes in scores between baseline and final-measurement (at3mat1). Contrary to the hypothesis, the increases in self-reported stress and negative affect were similar in C1:Paro_stress and C3:Bolster_stress, where participants had a later stress session. There was also no difference between C2:Stress_Paro and C4:Stress_bolster, where participants had an early stress session; both conditions showed similar responses in self-reported stress and negative affect. (Refer to Figure 5.11, page 215, and Figure 5.13, page 216).

Unexpected results were found in self-reported positive affect, where all conditions showed a decrease at the end of experiment when compared to baseline. (Refer to Figure 5.12, page 216).

Taking all the results together, based on the psychological responses, it can be concluded that the earlier session with either Paro or bolster did not help to reduce the stress and negative affect, and instead decreased the positive affect. Furthermore, none of the intervention sessions moderated the self-reported positive affect. But, the later session with Paro and bolster session did help to reduce the stress and negative affect. Hence, *H1(i)(a) is not supported*, and at the same time, *H1(ii)(a) is partially supported*.

The findings are in line with study by Lass-Hennemann et al. (2014), where the authors investigate the effects of being accompanied with either a dog, toy dog, unfamiliar person or just resting alone. The authors found a reduction in positive affect from baseline to the end of baseline. Lass-Hennemann et al. (2014) suggested that the reduction of positive affect was related to the stress induction and sample population used in their study. In Lass-Hennemann et al. (2014) study, the stressor used was a traumatic video showing a rape scene from a film, and most of the participants were female. Hence, the stressor might highly have affected the participants and the dog intervention did not moderate reduce the impact of the stressor. It might be that in the case of Study 3, *the stress induction was too stressful, hence reducing the impact of the intervention sessions (either Paro or bolster)*.

However, when compared to the recent study of Paro, the findings are in contrast with a study by Crossman et al. (2018). The authors found an increase in positive affect after interacting with Paro, even after participants had a stress session before the Paro session. A possible reason for this differences in results, is that *the stressor and sample population used in Study 3 was different from the study by* Crossman et al. (2018). In Crossman et al. (2018) study, the *participants recruited were aged between 6 and 9*, and the *stress induction used was a Trier Social Stress Test for Children*, where participants were asked to tell a story and complete a cognitive task in front of judges.

But, in Study 3, it worth noting that, there was a trend that showed a greater reduction in self-reported stress and negative affect in C2:Stress_Paro compared to C4:Stress_bolster. This may suggest there was a small effect of Paro intervention in reducing the effects of induced stress compared to bolster, even though the result is not significant. At the same time, there was a trend in both conditions of Paro that showed only a small reduction in positive affect compared to bolster conditions (See Figure 5.12, page 216). *This may suggest that the Paro conditions (regardless of whether participants had the Paro session before or after the stress session) have greater stress reducing effects compared to bolster conditions*. Despite the results not being significant, these findings are in line with the opinion of Khosla (2006) that positive affect plays an important role in helping people to cope and to minimize the stress reactivity during the stress exposure.

When considering the findings on psychological responses in all studies in this thesis, based on changes between baseline and post-measurement, the results of self-reported stress and positive affect are partially consistent across studies. While, *the results on self-reported negative affect are consistent across studies*.

In terms of self-reported stress and negative affect, these results suggest that decreased and increased negative affect depends on when the stress was induced (either before or after the intervention session). For example, if the stress session came after the intervention session, the negative affect increased. Then, when the intervention session came after the stress session, the negative affect decreased. These findings are in line with the opinion of Hamama et al. (2013) that changes in negative affect are more likely to be associated with exposure of stress.

However, in terms of self-reported positive affect, the findings may suggest that the changes in the results not only depend on when the stress was induced, but also depend on the types of stressor used. For example, in Study 1, the positive affect decreased after stress session, but the positive affect increased and returned to baseline after the Paro session. But, in Study 3 (in C2:Stress_Paro), the positive affect decreased after stress session, but the positive affect decreased further than baseline after the Paro session. Based on this pattern of findings, it suggests that the stressor used in Study 3 is probably more stressful than the stressor used in Study 1. This comparison was made as both has a similar intervention format.

5.10.2 Effects of conditions on physiological responses at the end of experiment compared to baseline

This section focuses on hypotheses H1(i)(b) and H1(ii)(b) and based on the results from physiological responses based on changes in scores between baseline and final-measurement (t4mt1).

In terms of physiological responses, all physiological responses (except HF) showed no effect of conditions, and all the physiological responses returned to baseline at the end of experiment. Hence *H1(i)(b) is not supported*, and *H1(ii)(b) is partially supported*.

As mentioned earlier, although all the physiological responses returned to baseline at the end of the experiment, there was an effect of conditions for HF. The HF was significantly higher in C2:Stress_Paro when compared to C1:Paro_stress and C3:Bolster_stress. There was also a trend that showed in C2:Stress_Paro, the HF was the highest among other three conditions. (Refer to Figure 5.31 on page 245). At the same time, the HF in C1:Paro_stress and C3:Bolster_stress showed a tendency to decrease from baseline to the end of experiment.

Then for other physiological responses, although the results are not significant between conditions, there was a trend in SCR that showed a small reduction in C2:Stress_Paro compared to other conditions. At the same time, a greater increase in SCR was observed in C4:Stress_bolster compared to other conditions (See Figure 5.28 on page 242). There was also a trend that showed a greater reduction in C1:Paro_stress and C2:Stress_Paro for HR (See Figure 5.29 on page 243). There was also a trend that the SDNN also showed a greater increase in C2:Stress_Paro compared to other conditions (See Figure 5.30 on page 244).

Taking all the results together, the increased HF and SDNN, and the decreased SCR and HR in C2:Stress_Paro suggests an increase in parasympathetic activity which could be related to excitement or relaxation (Park et al. 2018a). This indicate that the later intervention with Paro (in C2:Stress_Paro) might have had a small effect on reducing the effects of physiological stress responses compared to the later bolster session (in C4:Stress_bolster). Then, an increased in SCR and HR, and reduced SDNN in C4:Stress_bolster may suggest that participants were still affected by the previously induced stress. A further explanation about the changes in physiological responses will be discussed in next section 5.10.6, on page 253.

Facilitate faster recovery

Despite the non-significant difference between conditions on physiological responses from baseline to post-measurement, an effect of conditions was found in terms of accelerating a faster recovery. Based on changes in scores between stress session and post-measurement (t4mt3), a significant decrease in SCR in C2:Stress_Paro was found compared to bolster conditions (C3:Bolster_stress and C4:Stress_bolster), since C3:Bolster_stress and C4:Stress_bolster showed an increased trend in SCR. This suggests the later Paro condition may accelerate faster recovery in terms of SCR.

Then, in terms of HR, the subsequent bolster intervention also showed a small decrease in HR compared to other conditions, while the other conditions showed a greater decrease in HR. But, interestingly, there was also evidence that the earlier Paro and earlier bolster session showed a significant decrease from stress session to post-measurement in HR compared to the bolster condition (in C4:Stress_bolster), reflecting a faster recovery. *This suggests that the earlier intervention sessions might have impact on accelerating faster recovery of stress responses (in terms of HR and SCR).*

5.10.3 Effects of conditions during intervention session on psychological responses

This section focuses on hypotheses H2(a) and based on the results from psychological responses during intervention session.

No significant effect of conditions in self-reported stress and negative affect was found during the intervention session. All conditions showed low self-reported stress and negative affect. Only self-reported positive affect showed an effect of conditions. Taking all the results together, *H2(a) is not supported*.

Based on self-reported stress and negative affect, these findings are consistent with the findings in Study 2, where no differences were found in self-reported stress and negative affect scores were found, when comparing the scores of Paro condition and bolster condition.

However, as mentioned earlier, there was an effect of conditions for self-reported positive affect in Study 3. An earlier intervention with Paro (in C1:Paro_stress) showed significantly higher self-reported positive affect compared to other conditions. There was no significant difference between the later Paro condition (in C2:Stress_Paro) compared to the bolster sessions (in C3:Bolster_stress and C4:Stress_bolster). This may suggest that the effects of the earlier induced stress reduced the self-reported positive affect during the later intervention session in C2:Stress_Paro and C4:Stress_bolster. A possible reason for the different results between the Paro intervention sessions in C1:Paro_stress and C2:Stress_Paro, is that participants might have still been affected by the previously induced stress in C2:Stress_Paro, hence reducing the positive effects of Paro intervention. This finding is in line with the opinion of Frazier et al. (2004) that *the induced negative affect might have minimized the effectiveness of positive intervention.*

5.10.4 Effects of conditions during intervention session on physiological responses

This section focuses on hypotheses H2(b) and based on the results from physiological responses based on changes in scores between baseline and final-measurement (at3mat1).

No effect of conditions was found in HR and SDNN. Participants across conditions had similar levels of HR and SDNN either in earlier or later intervention sessions. But, there was an effect of conditions for SCR and HF. The SCR was significantly higher in C1:Paro_stress compared to C4:Stress_bolster, but there was no significant difference with other conditions. However, the HF was significantly higher in C2:Stress_Paro compared to the other three conditions. Therefore, *H2(b) is partially supported.*

As suggested from previous Study 1 and Study 2, the high SCR during Paro intervention was evoked by engagement with Paro. As suggested in the previous section, the high HF in C2:Stress_Paro is also related to engagement with Paro and is reflected as a coping skill that increased the parasympathetic activity during Paro session. At the same time, the bolster session (in C4:Stress_bolster) might also help in minimizing the impact of earlier stress responses.

5.10.5 Effects of conditions during the stress session

This section focuses on hypotheses H3(i)(a), H3(i)(b), H3(ii)(a) and H3(ii)(b) and based on the results from physiological responses based on changes in scores between baseline and final-measurement (t4mt1).

None of the hypotheses: H3 (i)(a, b) and H3 (ii)(a, b) are supported as no significant effect of conditions in stress session was found either both psychological and physiological responses.

This suggests that the stressor may not have been experientially different for participants.

The lack of a reduction in self-reported stress and negative affect in C1:Paro_stress compared to C3:Bolster_stress during the stress session, may suggest that both the earlier sessions (Paro and bolster) were not able to minimize the impact of stress during the stress session.

However, it is still too early to make a conclusion based on these results, therefore, in the next section, we will look at the results based changes of scores.

5.10.6 Intervention session versus stress session

In order to investigate further the effects of intervention and stress sessions on psychological and physiological responses, this section will discuss the results based on changes in scores:

(i) between baseline and Session 1 (at2mat1 or t2mt1), and

(ii) between Session 1 and Session 2 (at3mat2 or t3mt2).

Between baseline and Session 1

In terms of self-reported stress and negative affect, it was found that both the earlier intervention sessions significantly reduced the stress and negative affect, and that the earlier stress session significantly increased both measures. These findings provide evidence that the Paro and bolster session have a similar ability to reduce the effects of stress based on self-reported responses.

However, the Paro session induced significantly higher positive affect than the bolster session from baseline to Session 1. These results suggest that without any effects from previous session, when comparing the sessions to baseline, the Paro session can be considered as a positive intervention. This finding is in line with opinion of Sin & Lyubomirsky (2009) that mentioned that a positive intervention tended to induce positive feelings, and increase positive behaviours and thoughts.

In terms of physiological arousal, by comparing baseline and Session 1, it was found that both the intervention and the stress session significantly increased SCR and HR, and significantly decreased SDNN. There were no significant changes in HF.

These findings increase the understanding of the changes in physiological arousal during stress and positive intervention. There was no difference in arousal of SCR from baseline to Session 1 across conditions. This result is inconsistent with study by Mitsui et al. (2001*b*), where the authors found that the SCR was higher in stress session than in Paro session. However, the different results could be related to different stressors used in Study 3 and Mitsui et al. (2001*b*) study. Nevertheless, it worth noting a trend indicating that high arousal of SCR can be seen in Paro condition (C1:Paro_stress) compared to stress sessions (C2:Stress_Paro and C3:Bolster_stress) and bolster session (C4:Stress_bolster) (See Figure 5.21, page 232). This suggests that the positive intervention with Paro may be related to high arousal in SCR. This opinion is in line with the opinion of Balconi et al. (2009). Balconi et al. (2009) suggested a high arousal SCR is more likely to be found in pleasant stimuli compared to unpleasant stimuli.

In case of HR, the HR also increased significantly from baseline to Session 1 across conditions. At the same time, there was an effect of conditions, where the HR showed a significantly greater increase in stress sessions (C2:Stress_Paro and C4:Stress_bolster) compared to Paro (C1:Paro_stress) and bolster (C3:Bolster_stress) conditions. These findings suggest that both intervention sessions in C1:Paro_stress and C3:Bolster_stress showed a lower HR compared to stress session in C2:Stress_Paro and C4:Stress_bolster. These findings are consistent with Study 2, where participants interacting and stroking a bolster showed a low HR. But, at the same time, the findings in Study 3 are only partially in line with the study Mitsui et al. (2001*a*,*b*), where a high HR was found in the stress session (Mitsui et al. 2001*b*), a low HR was found in an active Paro session, but a high HR was found in a session where participants were asked to stroke a column covered with Paro's fur (Mitsui et al. 2001*a*).

Regarding the SDNN, all conditions decreased from baseline to Session 1. This finding is in contrast with the findings in Study 2, where the SDNN increased when interacting with and stroking the bolster compared to baseline. It was suggested in Study 2 that the increased SDNN was related to relaxation and excitement, as mentioned by Geisler et al. (2010), Park et al. (2018a). But, in Study 3, based on the observed trend, it can be seen that the SDNN showed a small decrease in C1:Paro_stress compared to other conditions (See Figure 5.30, page 244).

There are three possible explanations for the results of SDNN. The decreased SDNN could be related to (i) *the effects of stress induction*, (ii) *positive moods as a result of an engagement with Paro*, and (iii) *anticipation of the mathematical task*.

It was acknowledged that the acute stress such as that created by mathematical task usually reduces, but can also increase the SDNN (Castaldo et al. 2015, Rodrigues et al. 2018). In a review by Malik et al. (1996), it was claimed that the reduced SDNN during a stress task will also be accompanied with an increased HR. The elevated HR is usually related to an increase in sympathetic activity as response to the stress task. Therefore, in Study 3, based on increased HR, and increased self-reported stress and negative affect, and reduced self-reported positive affect, it may suggest that the reduced SDNN in C2:Stress_Paro and C4:Stress_bolster is related to stress induction. Then, there is a possibility that the reduced SDNN in C1:Paro_stress and C3:Bolster_stress was related to the positive reactions toward Paro, and effects of stroking a nice and comfortable object. This is reflected in both conditions (C1:Paro_stress and C3:Bolster_stress) which showed lower HR compared to C2:Stress_Paro and C4:Stress_bolster (where the participants had a stress session).

Furthermore, the reduced SDNN could be related to anticipation of a stressful situation, as suggested by Shin et al. (2016). Based on the math anxiety results, most of participants rated that they had high math anxiety across conditions, and participants might have been worried about the upcoming stress session in C1:Paro_stress and C3:Bolster_stress. Furthermore, although the results were not significant, but there was a trend that the SDNN showed a larger decreased in C3:Bolster_stress compared to C1:Paro_stress (See Figure 5.30 on page 244). This may suggest that the positive affect in C1:Paro_stress has a small effect in lessening the impact of anticipation to the stress session compared to C3:Bolster_stress. At the same time, these findings may support the important role of positive affect, in minimizing the effects of stress responses, in terms of physiological responses.

Between Session 1 and Session 2

Then, based on changes in scores between Session 1 and Session 2 (at3mat2), a similar direction of psychological changes was found between at3mat2 and at2mat1. The self-reported stress and negative affect increased significantly in the later stress session (for both condition C1:Paro_stress and C3:Bolster_stress), and there was an increased in positive affect in a later session with Paro (C2:Stress_Paro), but not in a later session with bolster (C4:Stress_bolster). However, the changes scores in positive affect between C2:Stress_Paro and C4:Stress_bolster is not significant.

Then, in terms of physiological responses based on changes in scores between Session 1 and Session 2 (t3mt2), the SCR increased in the Paro condition (C2:Stress_Paro), but reduced in the other three conditions. The HR was higher in both the later stress sessions, but reduced in the later intervention session (in C2:Stress_Paro and C4:Stress_bolster). There was also a trend that the HR in C3:Bolster_stress was higher than C1:Paro_stress, but the result was not significant. A similar trend in SCR and HR was also found between t3mt2 and t2mt1. Then, interestingly, there was a different trend in SDNN and HF based on changes in scores between Session 1 and Session 2. The HF and SDNN were higher in C2:Stress_Paro compared to other three conditions.

A possible explanation for this result is that *the increased SDNN and HF during Paro session in C2:Stress_Paro was related to increased parasympathetic activity, as a result from playing with Paro.* The findings are in line with the study by Shi et al. (2017) that found the increased SDNN related to positive emotions. These results are reflected in the Paro session only (in C2:Stress_Paro) that showed increased in self-reported positive affect from Session 1 to Session 2. This may suggest that the Paro intervention helps in reducing effects of previous stress in terms of physiological responses.

5.10.7 Effects of conditions on self-reported emotions

This section discusses emotions while playing with Paro or stroking a bolster.

Participants in C1:Paro_stress and C2:Stress_Paro rated significantly higher positive emotions compared to C3:Bolster_stress and C4:Stress_bolster, and all conditions rated low negative emotions This finding is in line with findings in Study 2, where participants also tended to rate higher positive emotions when playing with Paro compared to stroking a bolster. Based on these findings, *H2(a) for self-reported emotions is supported*. However, a manipulation of Paro session (either before or after stress induction) did not reduce the positive emotions towards Paro.

A similar trend on rated 'confused' emotion was also found in Paro conditions. Although the effects of condition were not statistically significant, but 'confused' was observed to be higher in C1:Paro_stress and C2:Stress_Paro. This finding suggests that participants tended to get confused in C1:Paro_stress and C2:Stress_Paro because that was the first time they met Paro and did not know how to interact or play with it, which is in line with the findings in Study 2 (in C2:no_familiar) In C2:no_familiar, participants were assigned to intervention with Paro without a chance to get a familiar with Paro before the experiment.

This finding is also in line with the study by Birks et al. (2016), where the authors mentioned that *participants tended to get confused as result of being unfamiliar with the Paro.* Birks et al. (2016) also mentioned that the lack of understanding about how interacting with Paro will benefit them, increased the negative perception towards it. However, in case of Study 3, interacting with Paro positively affected the participants, as reported by high self-reported positive affect in Paro conditions. The different result is possibly because a different sample population was used in the study by (Birks et al. 2016), that is elderly people with dementia.

5.10.8 Effects of conditions on observed facial expressions

The findings in Study 3 are in line with the findings in Study 1 and Study 2 on observed facial expression. Participants were showing significantly higher positive facial expressions in Paro sessions (C1:Paro_stress and C2:Stress_Paro) when compared to bolster sessions (C3:Bolster_stress and C4:Stress_bolster). Therefore, *H2(c) is supported.*

The findings in Study 3 suggested that participants were observed to show positive facial expressions when interacting with Paro, regardless of when they had been asked to interact with it, before or after the stress session. This also suggests that the previous stress induction did not affect their ways of interacting with Paro. This finding is in line with Study 2 (refer to section 4.10.3, page 139). Participants were observed smiling and laughing towards Paro robot compared to both bolster sessions, where participants tended to show bored expression (See Figure 5.18, page 226).

5.10.9 Effects of conditions on mathematical scores

H4 is about how the different intervention formats may have influenced participants' performance when doing mathematical task during the stress session.

A study by Allen et al. (2002) suggested that the presence of a pet dog when performing stressful task reduced the physiological reactivity to stress. Therefore, it was expected that participants would performed better in the mathematical task, as interacting with Paro showed a similar benefit to dog therapy, and hence should help to alleviate the anticipation of the stress session compared to other conditions. Unexpectedly, the results indicate that there was no significant difference between all conditions in mathematical task scores, and neither earlier interaction with Paro or bolster have an effect on the mathematical scores, therefore, *H4 is not supported.*

In further findings on mathematical scores between high scores and low scores group, no significant difference was found between the high and low scores categories in each condition. The results suggest that interacting with Paro and stroking a bolster might have not helped in improving mathematical performance during a stressful event, at least not in the context of short term interaction.

These findings are in line with a study by Trammell (2017), where the author found no difference in exam scores between participants that interacted with a dog and that did not interact with a dog. Trammell (2017) suggested that the dog might have reduced participants' levels of stress, but that the intervention may have a less impact on improving the exam scores in dog group. It was further mentioned by the author that the reason for no improvement was because the dog intervention was on the same day as the examination day (Trammell 2017). Therefore, it could be the case in Study 3, that the lack of effects of previous session with either Paro or bolster on mathematical scores was because they had to do the mathematical task immediately after the intervention session.

Another possible explanation for the lack of effects of intervention sessions on mathematical performance is math anxiety and noise annoyance. Based on a review by Foley et al. (2017), the authors mentioned that there is a relationship between high math anxiety and poor mathematical task performance. At the same time, Smith et al. (2010), Perham et al. (2013), Nasiri et al. (2014) also mentioned that the noise might affected the task performance. In Study 3, it was acknowledged that most participants rated higher in mathematical anxiety at the beginning of the study, and participants perceived the noise as annoyed, in all conditions. This may suggest that, *the high math anxiety and high annoyance to the noise influenced participants' performance in the mathematical task during the stress session.* However, *the short-term interventions may not have been sufficient to help in improving their mathematical performance or buffering the impact of noise exposure and math anxiety.*

In a further review by Foley et al. (2017), the authors mentioned that a positive intervention[6] immediately before performing the mathematical task, may temporarily reduce the math

[6]A positive intervention is defined as activities that helps to elevate positive feelings or moods (Foley et al. 2017).

anxiety. A study by (Park et al. 2014) supported that the positive intervention such as expression writing therapy help the participants with high in math anxiety to perform better. Despite the Paro intervention session (in C1:Paro_stress) in Study 3 considered as a positive intervention, as reflected in increased the self-reported positive affect, the findings are not in line with study by (Park et al. 2014). However, the contradictory results between studies might be related to different mathematical task questions and intervention used.

5.10.10 Effects of conditions on perceived stress

H5 is about how the different intervention formats may have influenced participants' perceived stress, based on a comparison between baseline (t1) and post-measurement (t4).

Study 3 found that there was no significant difference between baseline and post-measurement on self-reported perceived stress, therefore *H5 is not supported*. However, as can be seen in Figure 5.17, page 222, the perceived scores had a tendency to reduce in C2:Stress_Paro and C4:Stress_bolster. The trends may suggest that the later intervention sessions reduce the perceived stress.

In a further analysis, when comparing between low and high scores of perceived stress[7] groups, it was found that there was a significant increase in the number of participants that had lower perceived stress at the final measurement when compared to baseline in C1:Paro_stress. But, there were no significant changes in the low scores groups (as compared between baseline and post-measurement) in the other three conditions. Based on this finding, this may suggests that earlier interaction with Paro prior to stress session has a small effect on reducing the perceived stress in C1:Paro_stress.

This finding is in line with the opinion of Schiffrin & Nelson (2010), that there is a relationship between low perceived stress and positive emotion (such as being happy). In Study 3, as the intervention of Paro increased self-reported positive affect, it may have facilitated

[7]Perceived stress is defined as how an individual interprets a situation or an event as stressful (Folkman et al. 1986).

the reduction of self-reported perceived stress. However, no strong conclusion can be made, as only the participants in C1:Paro_stress showed decreased perceived stress (from baseline to post-measurement), and not participants in C2:Stress_Paro.

Nevertheless, these contradictory results can be explained by the differences in self-reported positive affect between C1:Paro_stress and C2:Stress_Paro. As mentioned in an earlier section, participants in C1:Paro_stress had significantly higher self-reported positive affect compared to C2:Stress_Paro during the intervention session. The high self-reported positive affect may help to reduce participants' perception of their stress. These findings are in line with the opinion of Lyubomirsky et al. (2005) who mentioned that high positive affect acts as moderator to help an individual to cope with stress. The lower self-reported positive affect in C2:Stress_Paro compared to C1:Paro_stress during the intervention session may be related to the previous stress session, hence reducing the effects of intervention. This in line with the opinion of Schiffrin & Nelson (2010), who mentioned that *low positive emotions may have affected the coping style, hence, no improvement was shown in perceived stress responses.*

5.11 Limitations in Study 3

The main limitation of Study 3 was that the timing of the collection of self-reported perceived stress was not controlled for all participants. All participants were asked to answer the self-reported perceived stress questionnaire prior to the experiment day, but each participant in C1:Paro_stress (for example) may have answered the questionnaire on a different day. But, the perceived stress asked how an individual perceived their stress within 1 month, and it was acknowledged that all participants answered the questionnaire at least within two weeks before the experiment day. The reason to collect participants' perceived stress before the experiment day was in order to examine participants' pre-existing stress before and after engaging in the experiment intervention. At the same time, the results may help to distinguish between pre-exiting stress and the effects of induced stress on the experiment day. In future, the study should account for the time when the participants need to answer the questionnaire, for example the experimenter can send an email to participant within 24 hours prior to the experiment day.

5.12 Conclusion

Overall, Study 3 presented two different intervention formats with a Paro or a bolster i.e. interaction with a Paro or a bolster before or after a stressor. The results from Study 3 suggest that the different intervention formats influenced the self-reported and physiological responses. At the same time, Study 3 also compared the immediate effect of Paro and bolster intervention, and its effects on reducing the effects of induced stress: environmental noise and mathematical task. The results suggest that both interventions showed a similar benefit to negative affect and self-reported stress, but that only the Paro intervention increased positive affect. In terms of reducing the effects of induced stress, it seems that neither of the interventions help to minimize the stress responses during the stress session, but the Paro intervention showed a positive improvement in physiological responses compared to the bolster intervention. At the same time, the factors (such as math anxiety, noise annoyance) that might have influenced the effects of intervention in reducing the stress responses were also discussed.

Chapter 6

Conclusion and future work

This chapter starts with outlining the aims and considering the main research questions of this thesis. Then limitations of this research are reviewed and several suggestions for future research are given.

6.1 The main aims of this thesis

There were three studies in this thesis. All three focused on young healthy adults or non-clinical population of students studying at University of Sheffield. Each study in this thesis is different from the others, however, all the studies share many common features. All the studies aimed to examine the effects of interacting with Paro, and the effectiveness of a Paro intervention on reducing stress responses.

Study 1 aimed to understand the effects of human and robot characteristics that influenced any effects that occurred during a short intervention session, and their influence on reducing the effects of induced stress (noise: traffic and noise conversation). Study 1 had four conditions: C1: a full interaction (i.e. talking and stroking) with an active Paro, C2: talking to an active Paro, C3: stroking an inactive Paro, and C4: no interaction with an inactive Paro. Label: C1:active_stroke_talk, C2:active_talk, C3:inactive_stroke, C4:inactive_rest.

Study 2 examined the effects of an earlier Paro intervention on reducing subsequent stress (noise: traffic, babies crying and noise conversation) responses, comparing them with the effects of a session with a bolster. At the same time, Study 2 also compared the effects of familiarity and non familiarity with Paro. This comparison was made in order to investigate the effects of novelty during intervention, and its role in reducing stress responses. Study 2 had four conditions: C1: get familiar with Paro before the experiment, then get another session with Paro, C2: a session with Paro without a chance to get familiar with it, C3: stroking a bolster, and C4: resting alone. Label: C1:familiar, C2:no_familiar, C3:bolster, C4:rest.

In the final study, Study 3 explored the effects of different intervention formats on induced stress (mathematical task and noise: traffic, babies crying and noise conversation). The intervention of Paro or bolster was either followed by a stress task or preceded by a stress task. Additionally, the study examined the effects of conditions on task performance scores and its effects on pre-existing stress. The pre-existing stress measured by Perceived Stress Scale (PSS) which asked about participants' pre-existing stress within one-month (Cohen et al. 1994). Study 3 had four conditions: C1: a session with Paro before the stress session, C2: a session with Paro after the stress session, C3: a session with bolster before the stress session, and C4: a session with bolster after the stress session. Label: C1:Paro_stress, C2:Stress_Paro, C3:Bolster_stress, C4:Stress_bolster

Table 6.1 summarized the study design of three studies in this thesis.

Table 6.1: Study design of three studies in this thesis.

Study	Study design	Intervention format	Label
1	C1: a full interaction (i.e. talking and stroking) with an active Paro, C2: talking to an active Paro, C3: stroking an inactive Paro, and C4: no interaction with an inactive Paro	Post-stress reducing	C1:active_stroke_talk, C2:active_talk, C3:inactive_stroke, C4:inactive_rest
2	C1: get familiar with Paro before the experiment, then get another session with Paro, C2: a session with Paro without a chance to get familiar with it, C3: stroking a bolster, and C4: resting alone	Pre-stress reducing	C1:familiar, C2:no_familiar, C3:bolster, C4:rest

Table 6.1 continued from previous page

Study	Study design	Intervention format	Label
3	C1: a session with Paro before the stress session, C2: a session with Paro after the stress session, C3: a session with bolster before the stress session, and C4: a session with bolster after the stress session	Pre- and post-stress reducing	C1:Paro_stress, C2:Stress_Paro, C3:Bolster_stress, C4:Stress_bolster

6.2 Main research questions in this thesis

There are four main research questions:

1. What are the immediate effects of short-term interaction with a Paro on observed facial expression, self-reported and physiological responses?
2. Does an intervention session with Paro have an impact on reducing effects of different kinds of induced stress on self-reported and physiological responses?
3. What effects does the implementation of different intervention formats (i.e. either interacting with Paro before or after stress inducement session) have on self-reported and physiological stress responses?
4. Does interacting with active Paro have a different impact on psychological and physiological responses than interacting with an inactive object i.e. an inactive Paro and a bolster?

In the next sections, a brief summary of the findings across three studies are given together with a discussion of the findings based on main research questions of this thesis.

6.3 Findings on research question 1

What are the immediate effects of short-term interaction with a Paro on observed facial expression, self-reported and physiological responses?

The effects of short-term interaction with Paro were measured using self-reported stress, positive and negative affect, and self-reported emotions, and using physiological responses (i.e. skin conductance response (SCR), heart rate (HR) and heart rate variability (HRV)). At the same time, participants interaction with Paro was observed based on facial expressions.

6.3.1 Self-reported responses

Across the three studies, the findings suggest that full interaction (an interaction that involves talking and stroking with Paro) with an active Paro resulted in high levels of momentary positive affect. Low self-reported stress and negative affect were also found across studies during the sessions with Paro. Similar results were found regardless of whether participants were asked to interact with either an active or inactive Paro (Study 1), or whether participants were familiar with the Paro or not (Study 2), or whether participants had a session with Paro after or before the stress session (Study 1 and Study 3). These results provide clear evidence that a short interaction with Paro is sufficient to induce high positive moods, low stress and low negative affect.

At the same time, the results highlighted that the physical embodiment of Paro such as its soft fur, movement, cute face and sound influenced the positive moods and stress relief. Engagement with the robot such as talking to or stroking it played a major role during the interaction with the robot.

6.3.2 Self-reported emotions and observed facial expressions

The findings on observed facial expressions during interactions with Paro were consistent across the three studies. Interacting with Paro results in a high frequency of positive facial expressions and low frequency of negative expressions. The findings on self-reported emotions (the self-reported emotions were not collected in Study 1) were also consistent across two studies (Study 2 and Study 3).

However, in Study 2, some effects of familiarisation with Paro before the experiment on self-reported emotions of 'confused' were found. This result highlights an understanding on the effects of novelty in social robot interaction. The results indicated that participants experienced some confusion when interacting with the Paro for a first time. This only occurred when the participants did not get a chance to become familiar with Paro before the experiment. Even though the intervention seemed to cause confusion but it did not influence participants' moods during the robot session. The findings showed that the self-reported positive affect was high and self-reported negative affect was low in both Paro conditions (in Study 2). These findings are in line with the opinion that even after the novelty effect of Paro has been removed, participants still tended to show an interest in playing with Paro (Sabanovic et al. 2013, Birks et al. 2016).

Interestingly, these effects were not only found in self-reported emotions but also observed in facial expressions. Based on observed facial expressions, there was a trend that showed that participants tended to laugh when meeting with Paro for the first time in C2:no_familiar although the difference to C1:familiar was not significant. While, participants in C1:familiar tended to show less laughing expression. The results were consistent with the opinion of Robinson et al. (2013c) that mentioned that participants in the healthy adult population tended to show excitement when interacting with active Paro.

6.3.3 Physiological responses

In terms of physiological responses, three physiological responses in total were recorded across studies: skin conductance response (SCR), heart rate (HR), and heart rate variability (HRV) parameters: Standard of all normal RR intervals (SDNN), and high-frequency (HF). There mixed results in physiological responses were found during the Paro session across the three studies.

Skin conductance responses

The increases in arousal of the skin conductance response (SCR) during interaction with Paro were consistent across the three studies. The arousal of SCR during Paro session is argued to have resulted from the positive emotions induced when interacting with the Paro, an argument supported by the self-reported positive affect and emotions across studies (Except for Study 1, where participants were no self-reported emotions were collected). High self-reported positive affect and emotions were found in all Paro conditions across the three studies.

Moreover, the findings in Study 3 also supported this argument. When comparing Paro session and noise sessions to baseline, both conditions showed high arousal of SCR, but the self-reported positive affect was also higher in Paro session compared to stress sessions. At the same time, the high arousal of SCR during Paro session also resulted from the physical movements (talking or stroking the Paro) that occurred during the interaction. The findings are in line with opinion that SCR are really sensitive to emotional changes (Figner & Murphy 2011), and at the same time the SCR is also sensitive to physical movement (Braithwaite et al. 2013).

Heart rate

The results of heart rate (HR) during the Paro session were consistent across two studies. These findings are in line with Mitsui et al. (2001*b*), where the HR showed no change from baseline to a session when participants interacted with active Paro.

Heart rate variability

In the Paro literature, only a study by Mitsui et al. (2001*b*) measured the effects of interacting with Paro using heart rate variability (HRV). The lack of studies of Paro using this measure could be because of the challenges in collecting and interpreting the data (Mitsui et al. 2001*a*, Kulic & Croft 2007, Sim et al. 2015). However, in the Mitsui et al. (2001*b*) study, the authors investigated a different HRV parameter i.e. ratio of low and high frequency. Therefore, this thesis is the first study that measures the effects of intervention with Paro using different HRV (the HRV was only collected in Study 1 and Study 2) parameters: i.e. standard of all normal RR intervals (SDNN) and high-frequency (HF).

The results of SDNN during Paro session were only partially consistent across two studies. In Study 2, when interacting with active Paro, the SDNN increased (from baseline to Paro session), and, in Study 3, the SDNN reduced. As discussed on page 254, possible explanations for this trend during interacting with Paro are related to (i) positive moods as a results of engagement with Paro, and (ii) anticipation of the mathematical task.

At the same time, the results of HF during Paro session were also partially consistent across two studies. There was no significant difference between Paro sessions in Study 2, in terms of HF. While, in Study 3, the HF was higher in the Paro condition (in C2:Stress_Paro: Participants had a Paro session after the stress session) compared to the Paro condition (in C1:Paro_stress: Participants had a stress session after the Paro session). A possible reason for these results may be related to recovery processes from the previous stress session (in C2:Stress_Paro). This suggests that interacting with Paro after the stress session helps in increasing the HF, which is related to parasympathetic activity (As discussed on page 255).

In general, the increased parasympathetic activity after stress exposure is important in order to regulate the autonomic nervous system (ANS) reactivity (Brown et al. 2013). The ability to cope with stress quickly not only helps to reduce the risk of cardiovascular diseases (such as heart attack (Wolff & Frishman 2004), but also helps to improve mental health (Ramanathan et al. 2017).

Conclusion Overall, the results of this thesis corroborated the findings of studies of short-term interaction with Paro such as Mitsui et al. (2001*b*), Crossman et al. (2018), where the Paro session was considered to be a positive intervention based on the improvement in positive moods after interacting with it. At the same time, this thesis also provide an insight into how interacting with Paro affected the physiological responses. These results are summarized in Table 6.2.

Table 6.2: Effects of conditions during intervention session. Notes: ns indicates no effect of conditions during the session; An asterisk indicates there was an effect of conditions during the session.

	Stress	PA	NA	SCR	HR	SDNN	HF
Study 1	*	ns	ns	*			
Study 2	ns	*	*	*	ns	ns	ns
Study 3	ns	*	ns	*	ns	ns	*

6.4 Findings on research question 2

Does an intervention session with Paro have an impact on reducing the effects of different kinds of induced stress on self-reported and physiological responses?

In research question 1, the results provide an insight into the effects of a short term session with Paro on self-reported and physiological responses. In order to address research question 2, a comparison between baseline and final measurement was made for each measure.

6.4.1 Effectiveness of the stressor

It was suggested that the stress induction methods used in this thesis significantly affected self-reported stress and negative affect throughout the three studies. This indicates that participants experienced mild stress during the stress induction sessions. However, does the Paro intervention reduce the effects of induced stress?

6.4.2 Self-reported responses

In Study 1, no changes were found in self-reported stress and positive affect (from baseline to post-measurement). However, the self-reported negative affect was significantly decreased (from baseline to post-measurement), although no effect of conditions was found. There was a non-significant trend suggesting that self-reported negative affect was lowest in the condition with full interaction with an active Paro compared to the other conditions. Nevertheless, Study 1 concluded that the presence of Paro in all conditions reduced the negative emotions (As discussed on page 99).

Then in Study 2, unexpectedly, no reduction (from baseline to post-measurement) was found in the self-reported responses to stress in any of the Paro conditions. The findings showed an increase in negative mood and stress, and a decrease in positive mood (from baseline to post-measurement).

Different results were found in Study 3, in terms of self-reported positive affect, where all conditions showed a decrease at the end of experiment when compared to baseline. The possible reasons for this may be related to the different intervention formats and different stressors used across studies (As discussed on page 185). It was also suggested that the stressor may have been too stressful in Study 2 and Study 3 (As discussed on page 171 and 247).

6.4.3 Physiological responses

In terms of physiological activity, this thesis provides an insight into how interacting with Paro affected participants' physiological stress responses.

In theory, as discussed on page 254, during relaxation or stress recovery heart rate variability (HRV), high-frequency (HF) and standard of all normal RR intervals (SDNN)) tends to increase, and whilst, heart rate (HR) tends to increase due to stress exposure. Specifically, skin conductance response (SCR) would tend to decrease or return to baseline when an individual is a in state of calm and able to control a distressing situation (Ulrich et al. 1991). By contrast, SCR would tend to increase due to engagement (Figner & Murphy 2011) or as a response to stress (Ulrich et al. 1991).

In this thesis, no changes in SCR were found across the studies (from baseline to post-measurement). Although, the SCR results were not significant, it is worth noting the trends suggested by SCR in this thesis. For example, in Study 2 (See figure 4.17 on page 144) and Study 3 (See figure 5.28 on page 242), at the end of experiment (compared to baseline), there was a decrease trend in SCR in the Paro condition compared to other conditions.

Additionally, in Study 3 (See figure 5.28 on page 242), despite the increased trend in SDNN (at the end of experiment compared to baseline) in C2:Stress_Paro (where participants had a subsequent session with Paro), at the same time, the SCR and HR decreased (See figure 5.29 on page 243), and the HF increased (See figure 5.31 on page 245), indicating that participants might have experienced some reduction in stress (as discussed on page 249).

However, the findings in Study 2 showed no improvement in stress reactivity in terms of HR and HRV in the Paro conditions. Interestingly, even though HR increased further from baseline at the end of experiment in all conditions, HR was significantly lower in the Paro conditions compared to resting alone condition. These findings may indicate that the Paro sessions helped in minimising the effect of subsequent stress on HR. (As discussed on page 173).

6.4.4 Perceived stress

Study 3, the thesis revealed significantly positive changes in pre-existing stress in the Paro condition (C1:Paro_stress). As discussed on page 259, these findings imply that a Paro intervention could help students cope with long-term stress. However, further research is needed to provide a strong conclusion, as only one of the Paro conditions that showed an effect on perceived pre-existing stress.

6.4.5 Mathematical performance

It was predicted that participants would perform better when they had an earlier Paro session before the stress session (in Study 3). Unexpectedly, no improvement in mathematical performance was found in the condition with an earlier interaction with Paro. This finding is not in line with study by Trammell (2017), where the authors found that participants in the dog condition performed better in their examination compared to the no dog condition. In the case of Study 3, the short intervention with Paro might not have been strong enough to reduce the impact of stress during the stress session, and to improve the performance of participants on the mathematical task (As discussed on page 258).

Conclusion Although the findings in most of the physiological responses were not statistically significant, interestingly, the results suggest that a later Paro session may helps to reduce the effects of stress in Study 1 and Study 3, but not in Study 2. These results are summarized in Table 6.3.

Table 6.3: Effects of time between baseline and post-measurement and the effects of conditions based changes scores from baseline to post-measurement. Notes: ns indicates no significant difference between baseline and post-measurement (t4mt1); increased indicates there was a significant increased from baseline to post-measurement; decreased indicates there was a significant decreased from baseline to post-measurement; mixed indicates there was a significant decreased and increased from baseline to post-measurement; An asterisk indicates there was an effect of conditions.

	Stress	PA	NA	SCR	HR	SDNN	HF
Study 1	ns	ns	decreased	ns			
Study 2	increased	decreased	increased	ns	increased*	increased*	increased
Study 3	mixed	decreased	mixed	ns	ns	ns	ns*

6.5 Findings on research question 3

What effect does the implementation of different intervention formats (i.e. interacting with Paro before or after stress inducement session) have on self-reported and physiological stress responses?

6.5.1 Self-reported responses

There was no difference in self-reported stress and negative affect, in terms of the different interventions formats across studies. However, the results reported in this thesis suggest that as well as intervention formats, the different types of stressor also have an effect on self-reported positive affect across studies (As discussed on page 247). For example, in Study 3, the positive affect decreased in all conditions at the end of experiment when compared to baseline. It was suggested that the decreased positive affect might relate to the stress session for being too stressful in Study 3.

At the same time, similar trends in psychological responses were observed during the stress session, across three studies. The trends are as follows: The self-reported negative affect and self-reported stress showed a significant increase from baseline to stress session (Study 1), a significant increase from intervention session to stress session (Study 2), and in Study 3, a significant increase in both intervention formats: (i) from baseline to stress session, and (ii) from intervention session to stress session. The self-reported positive affect was also low during the stress session across the three studies.

These findings are in line with the findings of McGlynn et al. (2016). In the Paro literature, only McGlynn et al. (2016) looked at the effects of stress before, during and after Paro was present. Although the findings in McGlynn et al. (2016) did not show statistical significant changes in self-reported stress, there was an increased trend of self-reported stress when Paro not present. Then when participants were resting alone in the presence of Paro, there was an indicative trend that showed a reduction in self-reported stress. Then, when Paro was removed again from the room, participants showed an increase in self-reported stress during the stressful situation.

6.5.2 Physiological responses

As discussed in the previous section 6.4.3, based on the results from heart rate variability (HRV), it was suggested the different intervention formats have a different impact on some physiological responses. For example, in Study 2, the standard of all normal RR intervals (SDNN) increased during intervention session, but in Study 3, the SDNN decreased during the intervention session.

Conclusion Overall, in terms of the different intervention formats and psychological and physiological responses, the Paro session may have a greater impact as *a stress reducing intervention* compared to as a preventive stress intervention. This conclusion is made based on the reduction in self-reported negative affect and stress (in Study 1 and Study 3), and improvement in parasympathetic activity (in Study 3), from baseline to post-measurement (in later Paro session). However, unexpectedly, the earlier Paro session did not have any effect on reducing the psychological stress responses to a subsequent stress session in Study 2 and Study 3.

6.6 Findings on research question 4

Does interacting with Paro have a different impact on psychological and physiological responses compared to stroking a bolster?

This thesis provides a considerable insight into the effects of an active Paro session as compared to an inactive object session such as inactive Paro (Study 1) or a bolster (Study 2 and Study 3).

6.6.1 Self-reported responses (During intervention session)

During the intervention session across studies, there are notable differences between the Paro session and bolster sessions, in terms of positive moods. In Study 1, participants in either active or inactive Paro conditions experienced high positive moods. While, in Study 2 and Study 3, the results showed that the Paro session induced more positive moods compared to bolster session. But, at the same time, in Study 2 and Study 3, both active Paro and bolster conditions provided a relaxing effect, as reflected in the results of low self-reported negative affect and stress.

6.6.2 Self-reported emotions and observed facial expressions

There are also notable differences between a Paro session and an inactive Paro (Study 1), and a bolster session (Study 2 and Study 3), in terms of self-reported emotions and observed facial expressions. Participants tended to rate positive emotions highly, and to show high positive observed facial expressions, in active Paro (across studies) compared to inactive Paro session (Study 1) and bolster session (Study 2 and Study 3).

6.6.3 Physiological responses (During intervention session)

During the intervention sessions across the studies, there are not many differences between the Paro session and the bolster session, in terms of physiological responses. In Study 1, both active or inactive Paro conditions induced high arousal in skin conductance response (SCR). Then, in Study 2, the results showed no difference in physiological responses between active Paro sessions and bolster session. The high arousal in SCR was also consistent across Study 1 and Study 2.

However, in Study 3, there are some differences in physiological responses between Paro and bolster sessions. A lower arousal in SCR was found in the later bolster session (in

C4:Stress_bolster) compared to other conditions. A possible reason for this may be related to stress reducing effect of stroking a nice object in C4:Stress_bolster (As discussed on page 251). At the same time, a later Paro session (in C2:Stress_Paro) showed a higher high-frequency (HF) compared to other conditions. A possible reason for this effect might be related to engagement with Paro in C2:Stress_Paro (As discussed on page 251). No difference was found in other physiological responses (i.e. heart rate (HR) and standard of all normal RR intervals (SDNN)) between Paro and bolster session.

6.6.4 Stress reducing effect of interventions

When comparing the effects of Paro and inactive objects, in terms of facilitating a stress reducing effect, most results showed no differences between conditions across studies.

In Study 1, most of the self-reported and physiological responses returned to baseline at the end of experiment. Interestingly, the self-reported negative affect decreased further from baseline at the end of experiment, in both active and inactive conditions. In Study 2, the results were different from Study 1. Both conditions, active and inactive Paro showed no reduction in self-reported and physiological stress responses. Then, in Study 3, the results for psychological and physiological responses are dependent on the intervention formats. No difference was found between active Paro and bolster conditions on psychological responses. There were also no differences between active Paro conditions and bolster conditions on physiological responses. Most of the physiological responses returned to baseline at the end of experiment in all conditions. At the same time, Study 3 found an improvement in physiological stress responses (in terms of high-frequency (HF)) in Paro conditions compared to bolster conditions. It was suggested that the improvement was dependent on intervention format (As discussed on page 261). As mentioned previously, this suggests that the Paro intervention is more effective as post-stress treatment than pre-stress treatment.

Furthermore, in Study 3, there was also plausible evidence that the Paro sessions facilitates faster recovery compared to bolster sessions. There were indicative trends in heart rate (HR), that showed the HR was elevated in C3:Bolster_stress, and that only showed a small decrease in C4:Stress_bolster, compared to Paro conditions, although the results are not significant between conditions.

Taken together, all these results (most of the results showed no difference between conditions: active Paro and inactive objects) are consistent with previous studies of Paro that compared the effects of Paro to inactive objects (such as an inactive Paro or a toy). A study by Moyle et al. (2017a) demonstrated that both conditions, active and inactive Paro, reduced the symptoms of agitation, increased pleasure and improved moods in elderly people with dementia. Then, when the effects of Paro were compared with a cat toy, the study by Thodberg et al. (2016) also did not find any differences between both conditions, both showed an improvement in depression scores.

Then, in terms of reaction towards active Paro and inactive Paro, a study by Takayanagi et al. (2014) observed a noticeable difference on reaction towards Paro and a lion toy. Participants tended to showed more positive facial expression towards Paro compared to a toy. The finding of Takayanagi et al. (2014) is consistent with the findings in this thesis, where participants were observed to show more positive reactions (based on observed facial expression and self-reported emotions collected during intervention session) towards active Paro than inactive objects.

Then in terms of the effects of Paro intervention on changes in physiological stress responses, only a study by Crossman et al. (2018) investigated the influence of Paro on reducing effects of induced stress. The findings in this thesis are partially consistent with study by Crossman et al. (2018), where the authors did not find any changes in hypothalamic pituitary adrenal (HPA) reactivity in both active and inactive Paro conditions.

Therefore, this thesis expands the literature of Paro on changes in physiological responses between interacting with active Paro and an inactive objects, and its potential on reducing the stress responses.

Conclusion Overall, based on the findings of the fourth research questions, between active Paro and inactive objects, it was suggested that the active Paro intervention has a greater beneficial effect on improving moods, and at the same time has some potential on reducing physiological stress responses compared to bolster intervention. These results are summarized in Table 6.4 and Table 6.5.

Table 6.4: Effects of conditions between active Paro and inactive objects during intervention session. Notes: ns indicates no effect of conditions between active Paro and inactive objects; An asterisk indicates there was an effect of conditions between active Paro and inactive objects.

	Stress	PA	NA	SCR	HR	SDNN	HF
Study 1	ns	ns	ns	ns			
Study 2	ns	*	ns	ns	ns	ns	ns
Study 3	ns	*	ns	*	ns	ns	*

Table 6.5: Effects of conditions between active Paro and inactive objects based on changes in scores between baseline and post-measurement. Notes: ns indicates no effect of conditions between active Paro and inactive objects; An asterisk indicates there was an effect of conditions between active Paro and inactive objects.

	Stress	PA	NA	SCR	HR	SDNN	HF
Study 1	ns	ns	ns	ns			
Study 2	ns	ns	ns	ns	ns	ns	ns
Study 3	*	ns	*	ns	ns	ns	*

6.7 Significant of this research

Based on the review on page 16 and 56, there are a large number of studies of Paro that found improvement in moods and reduced stress. However, most of the research focused on elderly people with dementia in nursing homes and hospitalised and autistic children (Pipitpukdee et al. 2011, Okita 2013, Robinson et al. 2013c). At the same time, only some studies investigated the effects of interacting with Paro on physiological responses (Mitsui et al. 2001*a*,*b*, Wada et al. 2005b, Crossman et al. 2018). Indeed, there are still many questions that need to be investigated in order to fully understand the potential benefits of social robot intervention for improving psychological and physiological health.

Possible reasons for the lack of studies investigating the effects of interacting with Paro on physiological responses are the challenges of collecting and interpreting physiological data. At the same time, the fact that physiological responses can be influenced by various factors that confound the results. However, physiological measurements have been widely used to assess the changes in autonomic nervous system (ANS), in order to monitor any risk associated with cardiovascular diseases. The findings in the thesis provide some considerable evidence of the changes in physiological responses as the result of interacting with Paro, of interacting with a nice comfortable object and as being exposed to an acute stress. The effects also were assessed using self-reported responses and observed video of facial expression to provide additional clues to any changes in physiological responses before, during and after the intervention with Paro.

The thesis expands the results of previous studies of Paro in several ways, as follows. There are no previous studies of Paro and physiological measures on non-clinical or healthy young adult populations using a large sample of participants. In terms of methodology, this thesis used random allocation of participants to conditions, comparing short-term interaction with Paro with a bolster, and controlled the independent variables and extraneous variables. In terms of research on induced stress and study in laboratory settings, to date there are no previous studies that look at the effects of interacting with Paro to acute stress such as *environmental noise and mathematical task*. The thesis also added new knowledge to existing literature on physiological responses to Paro and stress.

6.7.1 Sample size

A large sample size of participants (n = 208, the total is based on three studies) was used in this thesis as compared to previous studies of Paro that investigating the physiological improvement after interacting with Paro in healthy young adults (n = 6 to 10) (Mitsui et al. 2001*a*,*b*). However, this thesis did not provide the calculation for the power of sample size. But, some results from this thesis showed several significant values on psychological and physiological responses which implies the sample size was adequate.

6.7.2 Study design

This thesis used *a randomised controlled design and control conditions* that compared the effects of interacting with Paro with a bolster. The current studies also *carefully controlled the variables* that can potentially confound results when collecting physiological data. Potential confounding factors include age, gender, and the exclusion of participants with skin problems, heart problems and high blood pressure, mental health problems, hearing difficulties, and those using prescribed drugs medication.

Indeed, investigating the effects of Paro in a well-controlled environment can help to eliminate extraneous variables that may have confounded previous results. But, in this thesis, the findings found some improvement in moods and levels of stress as the result of short-term interactions with Paro. Therefore, this research could be further explored in natural settings such as in waiting room of dental clinic or emergency room, which commonly will induce acute stress. The induced stress can be related to the anticipation of meeting a doctor or received a treatment (Lundqvist et al. 2017, Santana et al. 2017), and may also related to noisy environment in the emergency room (Nahm et al. 2012).

Another suggestion, is that the Paro intervention could be implemented in universities. For example, at University of Sheffield, annually, during examination period, students have a chance to play with dogs in a Guide Dogs room (Wood et al. 2018). There was also a study that found a benefit of having a short session with dog therapy during examination periods (Trammell 2017). However, the short session with dog could be excluded to students who have fur allergies or are afraid of interacting with real animal, hence the session could be replaced by pet-robot therapy. Indeed, although, the findings from thesis found mixed results in different intervention formats (Paro as a pre- and post-treatment to stress), both Paro intervention sessions were able to induce momentary positive moods, and the positive moods seemed able to minimize and reduce the effects of acute stress.

6.7.3 Short-term positive intervention

This thesis provides more evidence for the Paro literature, in terms of short-term studies. This thesis also supports the claim the short duration sessions with Paro resulted in higher positive emotions and expressions compared to inactive objects. Participants rated themselves as higher in positive emotions after playing with Paro (either Paro being inactive or active). The research also found there was no difference in moods as a result of the novelty effects of the Paro intervention.The results reported here showed that participants continued to interact or play with Paro, and none of the participants gave up or refused to interact with Paro. Another interesting finding was that the participants reported that they felt less stress when stroking a bolster, but it seems that the bolster was more boring than fun.

6.7.4 Physiological responses

Previous studies of Paro suggested that interacting with Paro reduces blood pressure, heart rate, decreased stress responses by improving vital organs reactivity and improve moods (Wada & Shibata 2007b, Okita 2013, Robinson et al. 2013a,c). The thesis also supports these findings, which suggest improvement in heart rate variability (HRV) and a faster recovery in skin conductance response (SCR).

However, there were also partially consistent findings in this thesis to previous studies.

For example, in Study 2, interacting with Paro did not reduce the effects of induced stress. A previous study found that there was a reduction in heart rate (HR) after a 10 minutes' session with Paro (Robinson et al. 2013c). But it should be noted that the study by Robinson et al. (2013c) only focused on measuring the mere effects of interacting with Paro on heart rate, but not the effects after an immediate stress session.

The thesis also *added knowledge about physiological responses to the Paro literature, and stress literature*.

Even though, the thesis showed mixed results in HRV parameters across studies, the results help to provide an understanding of the changes in physiological responses during, before and after the interaction with Paro. At the same time, this thesis also provides an insight into the changes in physiological responses, from the results of interacting with the Paro, and from the results of exposure to stress.

Then, in terms of comparing the changes in physiological responses between Paro session and stress session, the results of this thesis were only partially consistent with Mitsui et al. (2001*b*). For example, Mitsui et al. (2001*b*) found that participants had higher physiological responses (SCR) in stress session compared to Paro session. But, in this thesis, the arousal of SCR were higher in the Paro session compared to the stress session (in Study 1 and Study 2), although in Study 3, a similar arousal of SCR was found between Paro session and stress session. As discussed on page 100 and 253, a possible reason for the different results may be related to the different stressors used in this thesis and in the study by Mitsui et al. (2001*b*).

Conclusion Overall, the results from this thesis suggests some potential benefits that interacting with Paro can have on cardiovascular activity, and how interacting with Paro may help to dampen the impact of acute stress on cardiovascular reactivity.

6.8 Limitations of this thesis

Although every effort has been made to design and implement good studies in this thesis, their limitations should be acknowledged. These limitations could be addressed in future research. As mentioned previously, the three studies were different from each other yet have many similar features. All the studies aimed to look at the effect of interacting with Paro and its effects on stress.

6.8.1 Heart rate variability

Firstly, there are some challenges when investigating and interpreting physiological responses such as heart rate variability (HRV). Two different sensors were used to derive the HRV data in this thesis. In Study 2, photoplethysmography (PPG) and in Study 3, electrocardiogram (ECG) was used for recording the HRV data. The PPG sensor was normally attached on fingertip; however, the ECG sensors were attached on participants' chests. The PPG derived the HRV data from arterial pulse and recorded using pulse oximetry. The pulse oximeter will produce a reflective light to the fingertip and the changes in light absorption based on increase of pulse pressure was measured by the PPG (Jeyhani et al. 2015). The ECG, however, measures the signal sent by the electrical activity of the heart (Eid et al. 2015). The use of two different sensors was due to availability of equipment that could be used by the researcher.

The PPG has been shown to be a good alternative to ECG because it is inexpensive and convenient to use to record HRV data for short and long-term (Lu et al. 2009). However, a limitation when recording using PPG, is that the PPG uses pulse-to pulse rather RR intervals (records by ECG) (Vescio et al. 2018). This could lead to problems when analysing the data, for example to determine the right peaks or waves, that is R peaks, which determine the frequency domain of HRV (Jeyhani et al. 2015).

Furthermore, when recording the PPG on fingertip, the data can produce some noises because of physical movement (Vescio et al. 2018). However, in Study 2, the PPG sensor was worn on the non-dominant hand and participants were instructed to avoid large movements of their non-dominant hand during the experiment.

ECG sensors are indeed a gold standard for measuring HRV parameters (Jeyhani et al. 2015). ECG also has some limitations such as data noise generated by the movement of the electrodes. The electrodes tend to dry after long term use and there was some electrical interference during recording also that should be noted (Davila et al. 2017).

Another challenge when measuring the physiological responses; previous research found that there was a significant difference in short-term HRV parameters based on age, gender and ethnicity. It was found that male and female participants aged above 50 showed lower parasympathetic activity (Voss et al. 2012). Research also suggested that females have lower HRV than males (Saleem et al. 2012). However, in this thesis, a gender balance in each study was not always achieved; only Study 1 managed to have a gender balance in each condition.

Additionally, the sympathetic activity is also lower in ethnically White Americans when comparing with African Americans (Choi et al. 2006). However, participants background such as ethnicity and courses were not recorded in Study 1 and Study 2, the information was only recorded in Study 3. Based on the report in Study 3, most participants were white and from a science and engineering background.

6.8.2 Stressors

Another limitation was that the induced stress could be too stressful and the short time interaction with Paro or bolster might not enough to reduce the stress. Indeed, in this thesis, the stressors used across studies were shown to be effective at inducing stress. However, the results from Study 3 showed a promising benefit of a 5 minute interaction with Paro on reducing daily acute stress.

6.8.3 Topic

Another limitation was that during the intervention session with active Paro, participants were asked to talk with Paro, but, the researcher did not record the conversation with Paro. This was because an aim of this thesis was to look at the emotions based on observed facial expressions. However, it would be interesting to assess the conversation between Paro and participants, in order to know more about the effects that conversation with Paro has on improving positive moods and reducing stress. At the same time, the researcher could suggest a topic to talk to Paro about to all the participants, in order to ensure the consistency of the conversation during the intervention.

6.8.4 Timed questionnaires

Lastly, some participants were not native English speakers, therefore some of them did not know the meaning of certain items, so they did not know how to rate their feeling according to the item. However, the researcher did provide a piece of paper that had the meaning of each item for them to refer to as they answered the questionnaires. A longer time spent answering the questionnaire between time-points might have affected the responses of the next session. This issue could be addressed by putting a time limit to answer each questionnaire. This is also could help to ensure participants did answer the questionnaire carefully, and would mean that the researcher could eliminate the data that could affect the results.

6.9 Future research

There a number of studies that can be done in the future, building on the work done in this thesis.

6.9.1 Reminiscence therapy

In this thesis, across the three studies, it has been observed that full interaction with Paro may have reduced the effects of stress responses. The full interaction with Paro involves stroking and talking. Indeed, stroking is one of the factors that contributed to stress reduction, as reflected in decreased stress after stroking either a Paro or a bolster. At the same time, it was also found that just talking to Paro also induced a positive mood and decreased stress in Study 1. Therefore, the potential benefit of this aspect should be further investigated, in order to enhance the effectiveness of short-term Paro intervention.

In the review in section 2.3.1 on page 19, it was shown that after interacting with Paro, elderly people tended to show an improvement in symptoms of depression. One of the improvements was that the elderly people started to talk with the Paro and other people in the nursing home (Moyle et al. 2013). It is commonly observed in people with dementia that they have a poor communication and that this can lead to social isolation (Moyle et al. 2013, Sung et al. 2014, Poscia et al. 2018). It was suggested that interaction with Paro may be able to remind elderly people of their past memory (Moyle et al. 2017b), which is related to the effects of reminiscence (Gallego-Perez et al. 2014).

Based on the Cambridge dictionary, reminiscence is defined as *"the act of remembering events and experiences from the past"* (Cambridge 2015). In terms of therapy, reminiscence helps to encourage social interaction among elderly people who have been diagnosed with depression (Tatchell et al. 2004). At the same time, study by Hallford & Mellor (2016) also found a positive effect of reminiscence therapy in reducing symptoms of depression in younger population.

In future research, the effects of Paro on reducing effects of acute stress could be further investigated by adopting a reminiscence-based therapy in the Paro intervention, as also suggested by Gallego-Perez et al. (2014). Participants will be asked to talk to Paro about a topic related to their past memories which made them happy. It would be anticipated that the reminiscence would influence the ways that participants interacted with Paro by increasings positive attitude towards it, hence increase the stress-reducing effect of Paro intervention on acute stress.

6.9.2 Hugging

Another aspect of human behaviour towards the robot is hugging. This thesis did not look at this effect, however, in previous studies of Paro, one of the behaviours observed in elderly people when interacting with Paro was hugging (for example, Shibata et al. (2009), Sabanovic et al. (2013)). However, these studies investigated the behaviour with stroking and hugging, hence the sole effects of hugging on improving moods or cardiovascular activity cannot be determined.

It was suggests that hugging helps to improve physiological responses to stress (Light et al. 2005, Gallace & Spence 2010, van Erp & Toet 2015), by increasing the levels of oxytocin, and regulate the blood pressure. At the same time, a review by Cohen et al. (2015) also suggested that the hugging is an important component in social support, and it has the potential to provide comfort and increased emotional experiences, hence benefiting social interaction. Future research could look at the effects on reducing acute stress, and whether hugging a Paro would have a greater impact on stress reactivity compare to talking or stroking.

6.10 Conclusion

In conclusion, this thesis provides an insight into the immediate effects of short-term interaction with Paro. The effects of interacting with active Paro were also compared to the effects of stroking an inactive object such as a bolster and inactive Paro. There are indicators that different types of stressor influenced the self-reported responses and physiological reactivity. In addition, this thesis also examines how the different intervention formats influenced the self-reported and physiological responses.

Importantly, this thesis has filled some of the gaps in the existing Paro literature through recruiting large samples of participants. At the same time, this thesis also expand the literature of Paro by conducting using different experimental design i.e. different types of stressor and manipulating the intervention formats. This thesis also not only to understand the effects of a Paro intervention on reducing, but also on protecting participants from stress responses. This thesis also has shown the different effects of intervention format on self-reported Positive and Negative Affect Schedule (PANAS) and perceived stress.

Overall, this thesis provides empirical evidence of the influence of short-term intervention with Paro on induced stress. This evidence should help other researchers to understand the factors that can influence psychophysiological responses to the Paro robot and to acute induced stress.

List of Acronyms

ANOVA Analysis Of Variance. 52, 117
ANS Autonomic Nervous System. 36, 46, 67, 268, 279
C1 Condition 1. 65
C2 Condition 2. 65
C3 Condition 3. 65
C4 Condition 4. 65
ECG Electrocardiogram. 24, 25, 42, 49, 179, 180, 283
EDR Electrodermal Response. 24, 25, 43
EEG Electroencephalography. 27
FDA Food And Drug Administration. 16
FNRIS Functional Near-infrared Spectroscopy. 25, 27
GAS General Adaptation Syndrome. 36
GSR Galvanic Skin Response. 43
HF High-frequency. 48, 109, 175, 182, 267, 268, 270, 276
HPA Hypothalamic Pituitary Adrenal. 59, 67, 277
HR Heart Rate. 7, 25, 42, 105, 109, 114, 161, 166, 169, 172, 174, 177, 179, 182, 185, 200, 245, 265, 267, 271, 276, 282
HRV Heart Rate Variability. 8, 25, 42, 48, 109, 114, 121, 161, 166, 169, 170, 172, 175, 177, 179, 182, 185, 200, 203, 245, 265, 267, 268, 270, 274, 281, 283
ICC Intraclass Correlation Coefficient. 79, 126, 224
MIT Montreal Imaging Task. 33
NA Negative Affect. 74
PA Positive Affect. 74
PANAS Positive And Negative Affect Schedule. 24, 34, 40, 74, 78, 121, 126, 197, 204, 288
PNS Parasympathetic Nervous System. 36
POMS Profile Of Mood State. 25, 40
PPG Photoplethysmography. 49, 120, 124, 179, 283
PSS Perceived Stress Scale. 41, 188, 263
RPM Raven's Progressive Matrices Test. 28, 58
SAM Self-Assessment Manikin. 34
SAT Separation Anxiety Test. 34
SC Skin Conductance. 42, 43, 121, 200
SCR Skin Conductance Response. 7, 66, 68, 69, 77, 92, 93, 94, 95, 97, 98, 99, 100, 101, 104, 105, 106, 121, 162, 169, 170, 171, 174, 177, 182, 185, 245, 265, 267, 271, 275, 281
SDNN Standard Of All Normal RR Intervals. 48, 109, 169, 172, 182, 267, 268, 270, 274, 276
SNS Sympathetic Nervous System. 36

SPSS Statistical Package For The Social Sciences. 54
SRI Stress Response Inventory. 41
SSID Student Services Information Desk. 72, 117, 319, 320, 321
STAI State Trait Anxiety Inventory. 24, 28, 34, 40
TSST Trier Social Stress Test. 33, 58
UMACL Mood Adjective Checklist. 40
UWIST University Of Wales Institute Of Science And Technology. 40

Bibliography

Abdi, J., Al-Hindawi, A., Ng, T. & Vizcaychipi, M. P. (2018), 'Scoping review on the use of socially assistive robot technology in elderly care', *BMJ open* **8**(2), e018815.

Allen, K., Blascovich, J. & Mendes, W. B. (2002), 'Cardiovascular reactivity and the presence of pets, friends, and spouses: The truth about cats and dogs', *Psychosomatic Medicine* **64**(5), 727–739.

Allen, K. M., Blascovich, J., Tomaka, J. & 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.

Alvarsson, J. J., Wiens, S. & Nilsson, M. E. (2010), 'Stress Recovery during Exposure to Nature Sound and Environmental Noise', *International Journal of Environmental Research and Public Health* **7**(3), 1036–1046.

Ambady, N. & Rosenthal, R. (1992), 'Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis', *Psychological Bulletin* **111**(2), 256–274.

American Psychological Association (2015), 'Understanding chronic stress'.
URL: http://www.apa.org/helpcenter/understanding-chronic-stress.aspx

Amiot, C. E. & Bastian, B. (2015), 'Toward a psychology of human–animal relations.', *Psychological Bulletin* **141**(1), 6.

Anderson, W. P., Reid, C. M. & Jennings, G. L. (1992), 'Pet ownership and risk factors for cardiovascular disease.', *The Medical journal of Australia* **157**(5), 298–301.

Andrews, B. & Wilding, J. M. (2004), 'The relation of depression and anxiety to life-stress and achievement in students.', *British journal of psychology (London, England : 1953)* **95**(Pt 4), 509–521.

Arkin, R. C. & Moshkina, L. (2014), 'Affect in human-robot interaction'.

Balconi, M., Brambilla, E. & Falbo, L. (2009), 'Appetitive vs. defensive responses to emotional cues. autonomic measures and brain oscillation modulation', *Brain research* **1296**, 72–84.

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

Barker, S. B., Knisely, J. S., McCain, N. L., Schubert, C. M. & Pandurangi, A. K. (2010), 'Exploratory study of stress-buffering response patterns from interaction with a therapy dog', *Anthrozoös* **23**(1), 79–91.

Barker, S. B., Rasmussen, K. G. & Best, A. M. (2003), 'Effect of aquariums on electroconvulsive therapy patients', *Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals* **16**(3), 229–240.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S. & Stansfeld, S. (2014), 'Auditory and non-auditory effects of noise on health', *The lancet* **383**(9925), 1325–1332.

Bassett, J. R., Marshall, P. M. & Spillane, R. (1987), 'The physiological measurement of acute stress (public speaking) in bank employees', *International Journal of Psychophysiology* **5**(4), 265–273.

Beck, A. M. (2014), 'The biology of the human-animal bond', *Animal Frontiers* **4**(3), 32–36.

Beck, A. M. & Katcher, A. H. (2003), 'Future directions in human-animal bond research', *American Behavioral Scientist* **47**(1), 79–93.

Beck, A. M. & Meyers, N. M. (1996), 'Health enhancement and companion animal ownership', *Annual review of public health* **17**(1), 247–257.

Beetz, A., Julius, H., Turner, D. & Kotrschal, K. (2012), 'Effects of social support by a dog on stress modulation in male children with insecure attachment', *Frontiers in Psychology* **3**(SEP), 1–9.

Beetz, A., Uvnäs-Moberg, K., Julius, H. & Kotrschal, K. (2012a), 'Psychosocial and psychophysiological effects of human-animal interactions: the possible role of oxytocin', *Frontiers in psychology* **3**, 234.

Bemelmans, R., Gelderblom, G. J., Jonker, P. & De Witte, L. (2012), 'Socially assistive robots in elderly care: A systematic review into effects and effectiveness', *Journal of the American Medical Directors Association* **13**(2), 114–120.

Bench, S. & Lench, H. (2013), 'On the Function of Boredom', *Behavioral Sciences* **3**(3), 459–472.

Berne, E. (1972), *What do you say after you say hello? : the psychology of human destiny*, Grove Press, Inc, New York.

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H. et al. (1997), 'Heart rate variability: origins, methods, and interpretive caveats', *Psychophysiology* **34**(6), 623–648.

Bethel, C. L. & Murphy, R. R. (2010), 'Review of Human Studies Methods in HRI and Recommendations', *International Journal of Social Robotics* **2**(4), 347–359.

Bharatharaj, J., Huang, L., Mohan, R., Al-Jumaily, A. & Krägeloh, C. (2017), 'Robot-Assisted Therapy for Learning and Social Interaction of Children with Autism Spectrum Disorder', *Robotics* **6**(1), 4.

Billman, G. E. (2011), 'Heart rate variability–a historical perspective', *Frontiers in physiology* **2**, 86.

BIOPAC (2017), 'Biopac systems'. [Online; accessed March 27, 2017].
URL: https://www.biopac.com/product/upgrade-to-mp36-system/

Birks, M., Bodak, M., Barlas, J., Harwood, J. & Pether, M. (2016), 'Robotic seals as therapeutic tools in an aged care facility: a qualitative study', *Journal of aging research* **2016**.

Bordens, K. S. & Abbott, B. B. (2002), *Research design and methods: A process approach*, 8 edn.

Boucsein, W. (1992), Electrodermal Activity as stress indicator, *in* 'Electrodermal Activity', Plenum Press, New York, chapter 3.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E. & Filion, D. L. (2012), 'Publication recommendations for electrodermal measurements', *Psychophysiology* **49**(8), 1017–1034.

Boxall, A. (2018), 'I had a one-night stand with a sleep robot'. [Online; accessed October 01, 2018].
URL: http://www.digitaltrends.com/cool-tech/somnox-sleep-robot-one-night-stand

Brace, N., Snelgar, R. & Kemp, R. (2016), *SPSS for Psychologists: And Everybody Else*, Macmillan International Higher Education.

Braithwaite, J. J., Watson, D. G., Jones, R. & Rowe, M. (2013), 'A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments', *Psychophysiology* **49**(1), 1017–1034.

Braun, C., Stangler, T., Narveson, J. & Pettingell, S. (2009), 'Animal-assisted therapy as a pain relief intervention for children', *Complementary Therapies in Clinical Practice* **15**(2), 105–109.

Brickel, C. M. (1979), 'The therapeutic roles of cat mascots with a hospital-based geriatric population: A staff survey', *Gerontologist* **19**(4), 368–372.

Broadbent, E. (2017), 'Interactions With Robots: The Truths We Reveal About Ourselves', *Annual Review of Psychology* **68**(1), 627–652.

Broadbent, E., Feerst, D. A., Lee, S. H., Robinson, H., Albo-Canals, J., Ahn, H. S. & MacDonald, B. A. (2018), 'How Could Companion Robots Be Useful in Rural Schools?', *International Journal of Social Robotics* .

Broadbent, E., Stafford, R. & MacDonald, B. (2009), 'Acceptance of healthcare robots for the older population: Review and future directions'.

Brodie, S. J., Biley, F. C. & Shewring, M. (2002), 'An exploration of the potential risks associated with using pet therapy in healthcare settings', *Journal of clinical nursing* **11**(4), 444–456.

Brougham, R. R., Zail, C. M., Mendoza, C. M. & Miller, J. R. (2009), 'Stress, sex differences, and coping strategies among college students', *Current Psychology* **28**, 85–97.

Brown, D. K., Barton, J. L. & Gladwell, V. F. (2013), 'Viewing nature scenes positively affects recovery of autonomic function following acute-mental stress', *Environmental science & technology* **47**(11), 5562–5569.

Buttelmann, D. & Römpke, A.-K. (2014), 'Anxiety-reducing effect: Dog, fish and plant in direct comparison', *Anthrozoös* **27**(2), 267–277.

Cable & Sensor (2010), '12-lead ecg placement guide with illustrations'. [Online; accessed June 30, 2017].
URL: https://www.cablesandsensors.com/pages/12-lead-ecg-placement-guide-with-illustrations

Cacioppo, J., Tassinary, L. G. & Berntson, G. G. (2007), *The Handbook of Psychophysiology*, Vol. 44, Cambridge University Press.

Calvo, M. G. & Nummenmaa, L. (2008), 'Detection of Emotional Faces: Salient Physical Features Guide Effective Visual Search', *Journal of Experimental Psychology: General* **137**(3), 471–494.

Calvo, M. G. & Nummenmaa, L. (2016), 'Perceptual and affective mechanisms in facial expression recognition: An integrative review', *Cognition and Emotion* **30**(6), 1081–1106.

Calvo, P., Fortuny, J. R., Guzmán, S., Macías, C., Bowen, J., García, M. L., Orejas, O., Molins, F., Tvarijonaviciute, A., Cerón, J. J. et al. (2016), 'Animal assisted therapy (aat) program as a useful adjunct to conventional psychosocial rehabilitation for patients with schizophrenia: results of a small-scale randomized controlled trial', *Frontiers in psychology* **7**, 631.

Calvo, R. a., Brown, I. & Scheding, S. (2009), 'Effect of experimental factors on the recognition of affective mental states through physiological measures', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **5866 LNAI**, 62–70.

Cambridge (2015), 'Cambridge Online Dictionary'.
URL: http://dictionary.cambridge.org

Capobianco, L., Morrison, A. P. & Wells, A. (2018), 'The effect of thought importance on stress responses: a test of the metacognitive model', *Stress* .

Carretero, R. & Sáez, P. (2017), Effects Of Non-pharmacological Intervention Through The Therapeutic Robot Paro In Persons With Alzheimer's Disease And Other Dementias, *in* 'Poster session presented at the Alzheimer's Global Summit Lisbon', number 18-19 September.

Casanova, G. M., Domanic, J., McCanne, T. R. & Milner, J. S. (1992), 'Physiological responses to non-child-related stressors in mothers at risk for child abuse', *Child Abuse & Neglect* **16**(1), 31–44.

Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M. & Pecchia, L. (2015), 'Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis'.

Center, R. (2015), 'Robot center robotic solutions provider'. [Online; accessed January 27, 2014].
URL: https://www.robotcenter.co.uk/products/copy-of-paro-therapeutic-robot-seal

Chandra, T., Yeates, D. B. & Wong, L. B. (2003), 'Heart rate variability analysis–current and future trends', *Business Briefing Global Health Care* **3**, 1–5.

Chang, S.-m. & Sung, H.-C. C. (2013), 'The effectiveness of seal-like robot therapy on mood and social interactions of older adults: a systematic review protocol', *JBI Database of Systematic Reviews and Implementation Reports* **11**(10), 68–75.

Charness, G., Gneezy, U. & Kuhn, M. A. (2012), 'Experimental methods: Between-subject and within-subject design', *Journal of Economic Behavior & Organization* **81**(1), 1–8.

Chatel-Goldman, J., Congedo, M., Jutten, C. & Schwartz, J.-L. (2014), 'Touch increases autonomic coupling between romantic partners', *Frontiers in behavioral neuroscience* **8**, 95.

Choi, J. B., Hong, S., Nelesen, R., Bardwell, W. A., Natarajan, L., Schubert, C. & Dimsdale, J. E. (2006), 'Age and ethnicity differences in short-term heart-rate variability', *Psychosomatic Medicine* **68**(3), 421–426.

Choi, S., Kim, S., Yang, J.-S., Lee, J.-H., Joo, C. & Jung, H.-I. (2014), 'Real-time measurement of human salivary cortisol for the assessment of psychological stress using a smartphone', *Sensing and Bio-Sensing Research* **2**, 8–11.

Civitci, A. (2015), 'The moderating role of positive and negative affect on the relationship between perceived social support and stress in college students.', *Educational Sciences: Theory and Practice* **15**(3), 565–573.

Coccia, C. & Darling, C. A. (2016), 'Having the Time of Their Life: College Student Stress, Dating and Satisfaction with Life', *Stress and Health* **32**(1), 28–35.

Coe, R. (2002), 'It's the effect size, stupid: What effect size is and why it is important'.

Cohen, S. & Janicki-Deverts, D. (2012), 'Who's stressed? distributions of psychological stress in the united states in probability samples from 1983, 2006, and 20091', *Journal of Applied Social Psychology* **42**(6), 1320–1334.

Cohen, S., Janicki-Deverts, D., Turner, R. B. & Doyle, W. J. (2015), 'Does hugging provide stress-buffering social support? a study of susceptibility to upper respiratory infection and illness', *Psychological science* **26**(2), 135–147.

Cohen, S., Kamarck, T. & Mermelstein, R. (1983), 'A Global Measure of Perceived Stress', *Journal of Health and Social Behavior* **24**(4), 385–396.

Cohen, S., Kamarck, T., Mermelstein, R. et al. (1994), 'Perceived stress scale', *Measuring stress: A guide for health and social scientists* pp. 235–283.

Cook, N. (2000), *Music : A very short introduction*, Oxford University Press, Oxford.

Cooper, D. (2018), 'Somnox's robotic pillow is designed to snuggle'.
URL: http://www.engadget.com/2018/01/09/somnox-s-robotic-pillow-is-designed-to-snuggle

Cowley, B., Filetti, M., Lukander, K., Torniainen, J., Henelius, A., Ahonen, L., Barral, O., Kosunen, I., Valtonen, T., Huotilainen, M. et al. (2016), 'The psychophysiology primer: a guide to methods and a broad review with a focus on human–computer interaction', *Foundations and Trends® in Human–Computer Interaction* **9**(3-4), 151–308.

Cramer, H., Kemper, N. a., Amin, A. & Evers, V. (2009), 'The effects of robot touch and proactive behaviour on perceptions of human-robot interactions.', *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09* pp. 275–276.

Crawford, J. R. & Henry, J. D. (2004), 'The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample.', *The British journal of clinical psychology / the British Psychological Society* **43**(Pt 3), 245–65.

Crossman, M. K., Kazdin, A. E. & Kitt, E. R. (2018), 'The influence of a socially assistive robot on mood, anxiety, and arousal in children', *Professional Psychology: Research and Practice* .

Cvijanović, N., Kechichian, P., Janse, K. & Kohlrausch, A. (2017), 'Effects of noise on arousal in a speech communication setting', *Speech Communication* **88**, 127–136.

Dang, T.-h.-h. H. H. & Tapus, A. (2013), 'Coping with stress using social robots as emotion-oriented tool: Potential factors discovered from stress game experiment', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8239 LNAI**, 160–169.

Darwin, C. (1872), 'The expression of the emotions in man and animals', *The American Journal of the Medical Sciences* **232**(4), 477.

Davila, M. I., Lewis, G. F. & Porges, S. W. (2017), 'The PhysioCam: A Novel Non-Contact Sensor to Measure Heart Rate Variability in Clinical and Field Applications', *Frontiers in Public Health* **5**.

Dawson, M. E., Schell, A. M. & Filion, D. L. (2007), 'The Electrodermal System'.

de Graaf, M. M. & Allouch, S. B. (2014), Evaluation of a socially assistive robot in eldercare, *in* 'Workshop an der International Conference on Human-Robot-Interaction'.

Dedovic, K., Renwick, R., Mahani, N. K., Engert, V., Lupien, S. J. & Pruessner, J. C. (2005), 'The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain', *Journal of Psychiatry and Neuroscience* **30**(5), 319–325.

Dehais, F., Sisbot, E. A., Alami, R. & Causse, M. (2011), 'Physiological and subjective evaluation of a human-robot object hand-over task', *Applied Ergonomics* **42**(6), 785–791.

Demello, L. R. (1999), 'The effect of the presence of a companion-animal on physiological changes following the termination of cognitive stressors', *Psychology and Health* **14**(5), 859–868.

DeSchriver, M. M. & Riddick, C. C. (1990), 'Effects of Watching Aquariums on Elders' Stress', *Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals* **4**(1), 44–48.

Dewing, J. (2009), 'Caring for people with dementia: noise and light.', *Nursing Older People* **21**(5).

Dhillon, B. & Liu, Y. (2006), 'Human error in maintenance: a review', *Journal of quality in maintenance engineering* **12**(1), 21–36.

Dickerson, S. S. & Kemeny, M. E. (2004), 'Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research', *Psychological bulletin* **130**(3), 355–391.

Dijkstra, K., Pieterse, M. E. & Pruyn, A. (2008), 'Stress-reducing effects of indoor plants in the built healthcare environment: The mediating role of perceived attractiveness', *Preventive medicine* **47**(3), 279–283.

Dirican, A. C. & Göktürk, M. (2011), 'Psychophysiological measures of human cognitive states applied in Human Computer Interaction', *Procedia Computer Science* **3**, 1361–1367.

Ditzen, B., Neumann, I. D., Bodenmann, G., von Dawans, B., Turner, R. a., Ehlert, U. & Heinrichs, M. (2007), 'Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women.', *Psychoneuroendocrinology* **32**(5), 565–74.

D'Mello, S. & Calvo, R. A. (2013), Beyond the basic emotions: what should affective computing compute?, *in* 'CHI'13 Extended Abstracts on Human Factors in Computing Systems', ACM, pp. 2287–2294.

Dong, J. G. (2016), 'The role of heart rate variability in sports physiology (Review)'.

Dotsinsky, I. (2007), 'Review of" advanced methods and tools for ecg data analysis", by gari d. clifford, francisco azuaje and patrick e. mcsharry (editors)', *Biomedical engineering online* **6**, 18.

Dursun, P., Emül, M., Gençöz, F. et al. (2010), A review of the literature on emotional facial expression and its nature., *in* 'Yeni Symposium', Vol. 48.

Dyrbye, L. N., Thomas, M. R. & Shanafelt, T. D. (2006), 'Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students.', *Academic Medicine* **81**(4), 354–373.

Eastwood, J. D., Frischen, A., Fenske, M. J. & Smilek, D. (2012), 'The unengaged mind: Defining boredom in terms of attention', *Perspectives on Psychological Science* **7**(5), 482–495.

Edwards, N. E. & Beck, A. M. (2002), 'Animal-assisted therapy and nutrition in Alzheimer's disease', *Western Journal of Nursing Research* **24**(6), 697–712.

Ehrén, E. (2014), How do educated therapy dogs (Canis lupus familiaris) get affected by their work in human care, PhD thesis, Bachelor's Thesis, University of Agricultural Sciences, Uppsala, Sweden.

Eid, D., Yousef, A. & Elrashidi, A. (2015), 'Ecg signal transmissions performance over wearable wireless sensor networks', *Procedia Computer Science* **65**, 412–421.

Eisenberg, D., Golberstein, E. & Gollust, S. E. (2007), 'Help-seeking and access to mental health care in a university student population', *Medical Care* **45**(7), 594–601.

Ekman, P. (1992), 'Are There Basic Emotions?', *Psychological Review* **99**(3), 550–553.

Ekman, P. (1993), 'Facial expression and emotion.', *The American psychologist* **48**(4), 384–392.

Ekman, P. (2003), *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*, New York.

Ekman, P., Friesen, W. V. & Hager, J. C. (2002), *Facial Action Coding System - Investigator's Guide*.

Etzi, R., Carta, C. & Gallace, A. (2018), 'Stroking and tapping the skin: behavioral and electrodermal effects', *Experimental brain research* **236**(2), 453–461.

FaceTurn (2011), 'FaceTurn The 3D Facial Expression Viewer'.
URL: http://www.faceturn.com/index.php

Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B. & Eastwood, J. D. (2013), 'Development and Validation of the Multidimensional State Boredom Scale', *Assessment* **20**(1), 68–85.

Fields, J. (1998), Recommendation for shared annoyance questions in noise annoyance surveys, *in* 'Proc. 7th International Congress on Noise as a Public Health Problem', Vol. 2, p. 481.

Figner, B. & Murphy, R. O. (2011), 'Using skin conductance in judgment and decision making research', *A handbook of process tracing methods for decision research* pp. 163–184.

Fine, A. H. (2010), *Handbook on Animal-Assisted Therapy*.

Fiocco, A. J. & Hunse, A. M. (2017), 'The buffer effect of therapy dog exposure on stress reactivity in undergraduate students', *International journal of environmental research and public health* **14**(7), 707.

Fish, M., Russoniello, C. & Clemmons-James, D. M. (2018), 'The use of biofeedback in college counseling', *Journal of College Counseling* **21**(3), 274–288.

Flannigan, G. & Dodman, N. H. (2001), 'Risk factors and behaviors associated with separation anxiety in dogs.', *Journal of the American Veterinary Medical Association* **219**(4), 460–466.

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C. & Beilock, S. L. (2017), 'The math anxiety-performance link: A global phenomenon', *Current Directions in Psychological Science* **26**(1), 52–58.

Folkman, S., Lazarus, R. S., Gruen, R. J. & DeLongis, A. (1986), 'Appraisal, Coping, Health Status, and Psychological Symptoms', *Journal of Personality and Social Psychology* **50**(3), 571–579.

Frazier, P. A., Tix, A. P. & Barron, K. E. (2004), 'Testing moderator and mediator effects in counseling psychology research.', *Journal of counseling psychology* **51**(1), 115.

Friedmann, E., Katcher, A. H., Lynch, J. J. & Thomas, S. A. (1980), 'Animal companions and one-year survival of patients after discharge from a coronary care unit.', *Public Health Reports* **95**(4), 307–312.

Friedmann, E. & Thomas, S. A. (1995), 'Pet ownership, social support, and one-year survival after acute myocardial infarction in the Cardiac Arrhythmia Suppression Trial (CAST)', *The American Journal of Cardiology* **76**(17), 1213–1217.

Frodi, A. M., Lamb, M. E., Leavitt, L. A. & Donovan, W. L. (1978), 'Fathers' and mothers' responses to infant smiles and cries', *Infant Behavior and Development* **1**, 187–198.

Fukuda, S., Morimoto, K., Lifestyle, ., Fukuda, S. & Morimoto, K. (2001), 'Lifestyle, Stress and Cortisol Response: Review II', *Environmental Health and Preventive Medicine* **6**(1), 15–21.

Gallace, A. & Spence, C. (2010), 'The science of interpersonal touch: An overview', *Neurosci. Biobehav. Rev.* **34**(2), 246–259.

Gallego-Perez, J., Lohse, M. & Evers, V. (2014), Robots for the psychological wellbeing of the elderly, *in* 'HRI 2014 Workshop Socially Assistive Robots for the Aging Population: Are we trapped in stereotypes'.

Geen, R. G. & McCown, E. J. (1984), 'Effects of noise and attack on aggression and physiological arousal', *Motivation and Emotion* **8**(3), 231–241.

Geisler, F. C., Vennewald, N., Kubiak, T. & Weber, H. (2010), 'The impact of heart rate variability on subjective well-being is mediated by emotion regulation', *Personality and Individual Differences* **49**(7), 723–728.

Gelderblom, G. J., Bemelmans, R., Spierts, N., Jonker, P. & De Witte, L. (2010), Development of paro interventions for dementia patients in dutch psycho-geriatric care, *in* 'International Conference on Social Robotics', Springer, pp. 253–258.

Ghaddar, A., Omar, K. H., Dokmak, M., Kansour, N. A., Jbara, Z., Laham, S. & Ali, S. (2013), 'Work-related Stress and Urinary Catecholamines among Laboratory Technicians', *Journal of Occupational Health* **55**(5), 398–404.

Ghasemi, A. & Zahediasl, S. (2012), 'Normality tests for statistical analysis: a guide for non-statisticians', *International journal of endocrinology and metabolism* **10**(2), 486.

Gidlof-Gunnarsson, A., Ohrström, E. & Ogren, M. (2007), 'Noise annoyance and restoration in different courtyard settings: Laboratory experiments on audio-visual interactions', *Proceedings of the 2007 International Congress and Exhibition on Noise Control Engineering* (August).

Goedhart, A. D., Willemsen, G., Houtveen, J. H., Boomsma, D. I. & De Geus, E. J. (2008), 'Comparing low frequency heart rate variability and preejection period: Two sides of a different coin', *Psychophysiology* **45**(6), 1086–1090.

Goines, L. & Hagler, L. (2007), 'Noise pollution: a modem plague.', *The Southern medical journal* **100**(3), 287–294.

Goldstein, D. S. (2010), 'Adrenal responses to stress', *Cellular and Molecular Neurobiology* **30**(8), 1433–1440.

Greer, B. & Mulhern, G. (2001), *Making Sense of Data and Statistics in Psychology*, Macmillan Education UK.

Groh, A. M. & Roisman, G. I. (2009), 'Adults' autonomic and subjective emotional responses to infant vocalizations: The role of secure base script knowledge.', *Developmental psychology* **45**(3), 889.

Gunawardena, N. (2011), 'Choosing the correct statistical test in research', *Sri Lanka Journal of Child Health* **40**(4).

Gustafsson, C., Svanberg, C. & Mullersdorf, M. (2015), 'Using a Robotic Cat in Dementia Care A Pilot Study', *Journal of Gerontological Nursing* **41**(10), 46–56.

Hallford, D. J. & Mellor, D. (2016), 'Autobiographical memory-based intervention for depressive symptoms in young adults: A randomized controlled trial of cognitive-reminiscence therapy', *Psychotherapy and psychosomatics* **85**(4), 246–249.

Hamama, L., Ronen, T., Shachar, K. & Rosenbaum, M. (2013), 'Links between stress, positive and negative affect, and life satisfaction among teachers in special education schools', *Journal of Happiness Studies* **14**(3), 731–751.

Healey, J. F. (2014), *Statistics: A tool for social research*, Cengage Learning.

Heath, S. & Kenyon, L. (2001), 'Single Young Professionals and Shared Household Living', *Journal of Youth Studies* **4**(1), 83–100.

Heimlich, K. (2001), 'Animal-assisted therapy and the severely disabled child: A quantitative study', *Journal of Rehabilitation* **67**, 48–54.

Henelius, A., Hirvonen, K., Holm, A., Korpela, J. & Müller, K. (2009), 'Mental workload classification using heart rate metrics', *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009* **2009**, 1836–1839.

Herzog, H. (2011), 'The impact of pets on human health and psychological well-being: fact, fiction, or hypothesis?', *Current Directions in Psychological Science* **20**(4), 236–239.

Hirshfield, L., Chauncey, K., Gulotta, R., Girouard, A., Solovey, E., Jacob, R., Sassaroli, A. & Fantini, S. (2009), Combining electroencephalograph and near infrared spectroscopy to explore users' instantaneous and continuous mental workload states, *in* 'HCI International', Vol. 2009.

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U. & Søgaard, K. (2004), 'The effect of mental stress on heart rate variability and blood pressure during computer work', *European Journal of Applied Physiology* **92**(1-2), 84–89.

Hoffmann, M., Heinke, S., Lang, K., Leder, U., Hoffmann, G., Heinke, M., Dannberg, G., Kuhnert, H. & Muller, S. (1996), Comparative studies of heart rate variability (HRV) using Holter-ECG-systems and high resolution ECG-systems in patients with acute myocardial infarction and dilatative cardiomyopathy and healthy volunteers, Vol. 3, pp. 1256–1257.

Holland, P. (2018), 'Fur-covered robots are the new therapy animal'.
URL: http://www.cnet.com/news/comfort-robots-aflac-duck-and-paro-seal-help-sick-people

Hooker, S. D., Freeman, L. H. & Stewart, P. (2002), 'Pet therapy research: a historical review.', *Holistic nursing practice* **16**(5), 17–23.

Horsten, M., Ericson, M., Perski, A., Wamala, S. P., Schenck-Gustafsson, K. & Orth-Gomer, K. (1999), 'Psychosocial factors and heart rate variability in healthy women', *Psychosom Med* **61**(1), 49–57.

Howitt, D. (2017), *First Steps in Research and Statistics: A Practical Workbook for Psychology Students*, Routledge.

Hsu, S.-M., Ko, W.-J., Liao, W.-C., Huang, S.-J., Chen, R. J., Li, C.-Y. & Hwang, S.-L. (2010), 'Associations of exposure to noise with physiological and psychological outcomes among post-cardiac surgery patients in ICUs.', *Clinics* **65**(10), 985–989.

Hudd, S. S., Dumlao, J., Erdmann-Sager, D., Murray, D., Phan, E., Soukas, N. & Yokozuka, N. (2000), 'Stress at college: Effects on health habits, health status and self-esteem', *College Student Journal* **34**(3560), 217–228.

Hunt, J. & Eisenberg, D. (2010), 'Mental health problems and help-seeking behavior among college students', *Journal of Adolescent Health* **46**(1), 3–10.

Iannuzzi, D. & Rowan, A. N. (1991), 'Ethical issues in animal-assisted therapy programs', *Anthrozoös* **4**(3), 154–163.

Iffland, B., Sansen, L. M., Catani, C. & Neuner, F. (2014), 'Rapid heartbeat, but dry palms: reactions of heart rate and skin conductance levels to social rejection', *Frontiers in Psychology* **5**(August), 1–10.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R. & Schyns, P. G. (2012), 'Facial expressions of emotion are not culturally universal', *Proceedings of the National Academy of Sciences* **109**(19), 7241–7244.

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. & Schyns, P. G. (2016), 'Four not six: Revealing culturally common facial expressions of emotion', *Journal of Experimental Psychology: General* **145**(6), 708–730.

Janipha, N. A. I., Mustapha, A. A. & Ismail, F. (2012), 'Workplace Stress amongst Consultants in Practice', *Procedia - Social and Behavioral Sciences* **68**, 183–191.

Jarolmen, J. & Patel, G. (2018), 'The effects of animal-assisted activities on college students before and after a final exam', *Journal of Creativity in Mental Health* pp. 1–11.

Jeyhani, V., Mahdiani, S., Peltokangas, M. & Vehkaoja, A. (2015), Comparison of hrv parameters derived from photoplethysmography and electrocardiography signals, *in* 'Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE', IEEE, pp. 5952–5955.

Johnston-Wilder, S., Brindley, J. & Dent, P. (2014), 'A survey of mathematics anxiety and mathematical resilience among existing apprentices'.

Joranson, N., Pedersen, I., Rokstad, A. M. M. & Ihlebæk, C. (2015), 'Effects on Symptoms of Agitation and Depression in Persons With Dementia Participating in Robot-Assisted Activity: A Cluster-Randomized Controlled Trial', *Journal of the American Medical Directors Association* **16**(10), 867–873.

Jordan, T. R., Khubchandani, J. & Wiblishauser, M. (2016), 'The Impact of Perceived Stress and Coping Adequacy on the Health of Nurses: A Pilot Investigation', *Nursing Research and Practice* .

Jung, M. M., van der Leij, L. & Kelders, S. M. (2017), 'An exploration of the benefits of an animallike robot companion with more advanced touch interaction capabilities for dementia care', *Frontiers in ICT* **4**, 16.

Kachouie, R., Sedighadeli, S., Khosla, R. & Chu, M.-T. (2014), 'Socially Assistive Robots in Elderly Care: A Mixed-Method Systematic Literature Review', *International Journal of Human-Computer Interaction* **30**(5), 369–393.

Karim, N., Hasan, J. A. & Ali, S. S. (2011), 'Heart rate variability: a review', *Journal of Basic and Applied Sciences* **7**(1), 71–77.

Kawaguchi, Y., Wada, K., Okamoto, M., Tsujii, T., Shibata, T. & Sakatani, K. (2012), 'Investigation of brain activity after interaction with seal robot measured by fNIRS', *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication* pp. 571–576.

Kennedy, A. (2015), 'Impact of Emotional Support Animals on Student Stress'.

Khosla, M. (2006), 'Positive affect and coping with stress', *Journal of the Indian Academy of Applied Psychology* **32**(3), 185–192.

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

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H. & Koo, B.-H. (2018), 'Stress and heart rate variability: A meta-analysis and review of the literature', *Psychiatry investigation* **15**(3), 235.

Kjellberg, A., LandstrÖM, U., Tesarz, M., Söderberg, L. & Åkerlund, E. (1996), 'The effects of nonphysical noise characteristics, ongoing task and noise sensitivity on annoyance and distraction due to noise at work', *Journal of Environmental Psychology* **16**(2), 123–136.

Klein, B. & Cook, G. (2012), Emotional robotics in elder care - A comparison of findings in the UK and Germany, *in* 'Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)', Vol. 7621 LNAI, Elsevier Ltd, pp. 108–117.

Kobayashi, A., Yamaguchi, Y., Ohtani, N. & Ohta, M. (2017), 'The Effects of Touching and Stroking a Cat on the Inferior Frontal Gyrus in People', *Anthrozoos* **30**(3), 473–486.

Kogan, L. R., Hellyer, P. W., Clapp, T. R., Suchman, E., McLean, J. & Schoenfeld-Tacher, R. (2018), 'Use of short animal-themed videos to enhance veterinary students' mood, attention, and understanding of pharmacology lectures', *Journal of veterinary medical education* **45**(2), 188–194.

Koh, K. B., Park, J. K., Kim, C. H. & Cho, S. (2001), 'Development of the stress response inventory and its application in clinical practice.', *Psychosomatic medicine* **63**(4), 668–78.

Kolling, T., Baisch, S., Schall, A., Selic, S., Rühl, S., Kim, Z., Rossberg, H., Klein, B., Pantel, J., Oswald, F. & Knopf, M. (2016), Chapter 5 - what is emotional about emotional robotics?, *in* 'Emotions, Technology, and Health', Emotions and Technology, Academic Press, pp. 85 – 103.

Kramer, S. C., Friedmann, E. & Bernstein, P. L. (2009), 'Comparison of the effect of human interaction, animal-assisted therapy, and aibo-assisted therapy on long-term care residents with dementia', *Anthrozoös* **22**(1), 43–57.

Kroes, S. (2005), 'Detecting boredom in meetings', *University of Twente* pp. 1–5.

Krypotos, A.-M., Jahfari, S., van Ast, V. A., Kindt, M. & Forstmann, B. U. (2011), 'Individual Differences in Heart Rate Variability Predict the Degree of Slowing during Response Inhibition and Initiation in the Presence of Emotional Stimuli', *Frontiers in Psychology* **2**, 278.

Kulic, D. & Croft, E. (2007), 'Physiological and subjective responses to articulated robot motion', *Robotica* **25**(01), 13.

Kumar, S., Kumar, D. K., Alemu, M. & Burry, M. (2004), Detection of the Human Stress Response to Auditory Noise, *in* 'Proceedings of Accoustics', pp. 599–602.

Kumaraswamy, N. (2013), 'Academic stress, anxiety and depression among college students-A brief review', *International Review of Social Sciences and Humanities* **5**(1), 135–143.

Kushki, A., Fairley, J., Merja, S., King, G. & Chau, T. (2011), 'Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites', *Physiological Measurement* **32**(10), 1529–1539.

Lajante, M., Droulers, O., Dondaine, T. & Amarantini, D. (2012), 'Opening the "black box" of electrodermal activity in consumer neuroscience research.', *Journal of Neuroscience, Psychology, and Economics* **5**(4), 238.

Lakey, B. & Cohen, S. (2000), 'Social support and theory', *Social support measurement and intervention: A guide for health and social scientists* pp. 29–52.

Landers, R. (2015), ‘Computing (ICC) as Intraclass Estimates Correlations of Interrater Reliability in SPSS’, *The Winnower* (2:e143518.81744 (2015)), 1–4.

Lang, P. J. (1995), ‘The Emotion Probe: Studies of Motivation and Attention’, *American Psychologist* **50**(5), 372–385.

Lass-Hennemann, J., Peyk, P., Streb, M., Holz, E. & Michael, T. (2014), ‘Presence of a dog reduces subjective but not physiological stress responses to an analog trauma.’, *Frontiers in Psychology* **5**(September), 1–7.

Leger, E. (1996), Mature students’ perception of stress on returning to learn, PhD thesis, University of Leeds.

Leite, I., Martinho, C. & Paiva, A. (2013), ‘Social robots for long-term interaction: a survey’, *International Journal of Social Robotics* **5**(2), 291–308.

Leroy, G. & Miller, T. (2010), ‘Perils of providing visual health information overviews for consumers with low health literacy or high stress’, *Journal of the American Medical Informatics Association* **17**(2), 220–223.

Levinson, B. M. (1962), ‘The dog as a ”co-therapist”’, *Mental Hygiene* **46**(1), 59–65.

Levsen, N. (2015), *Lead Markets in Age-Based Innovations: Demographic Change and Internationally Successful Innovations*, Forschungs-/Entwicklungs-/Innovations-Management, Springer Fachmedien Wiesbaden.

Liang, A., Piroth, I., Robinson, H., MacDonald, B., Fisher, M., Nater, U. M., Skoluda, N. & Broadbent, E. (2017), ‘A Pilot Randomized Trial of a Companion Robot for People With Dementia Living in the Community’, *Journal of the American Medical Directors Association* **18**(10), 871–878.

Libby, D. J., Worhunsky, P. D., Pilver, C. E. & Brewer, J. A. (2012), ‘Meditation-induced changes in high-frequency heart rate variability predict smoking outcomes’, *Frontiers in human neuroscience* **6**, 54.

Light, K. C., Grewen, K. M. & Amico, J. A. (2005), ‘More frequent partner hugs and higher oxytocin levels are linked to lower blood pressure and heart rate in premenopausal women’, *Biological Psychology* **69**(1), 5–21.

Lindvall, T. & Radford, E. P. (1973), ‘Measurement of annoyance due to exposure to environmental factors: The fourth karolinska institute symposium on environmental health’, *Environmental research* **6**(1), 1–36.

Loeb, M. (1986), *Noise and human efficiency*, John Wiley & Sons.

London, H., Schubert, D. S. & Washburn, D. (1972), ‘Increase of autonomic arousal by boredom.’, *Journal of Abnormal Psychology* **80**(1), 29.

Lottridge, D., Chignell, M. & Jovicic, A. (2011), ‘Affective Interaction: Understanding, Evaluating, and Designing for Human Emotion’, *Reviews of Human Factors and Ergonomics* **7**(1), 197–217.

Lovell, B., Moss, M. & Wetherell, M. A. (2011), 'Perceived stress, common health complaints and diurnal patterns of cortisol secretion in young, otherwise healthy individuals', *Hormones and behavior* **60**(3), 301–305.

Lu, G., Yang, F., Taylor, J. A. & Stein, J. F. (2009), 'A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects', *Journal of Medical Engineering and Technology* **33**(8), 634–641.

Lu, S., Zhao, H., Ju, K., Shin, K., Lee, M., Shelley, K. & Chon, K. H. (2008), 'Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information?', *Journal of Clinical Monitoring and Computing* **22**(1), 23–29.

Luijcks, R., Hermens, H. J., Bodar, L., Vossen, C. J., Van Os, J. & Lousberg, R. (2014), 'Experimentally induced stress validated by EMG activity', *PLoS ONE* **9**(4), 1–8.

Lumley, M. A. & Provenzano, K. M. (2003), 'Stress management through written emotional disclosure improves academic performance among college students with physical symptoms.', *Journal of Educational Psychology* **95**(3), 641–649.

Lund, H. G., Reider, B. D., Whiting, A. B. & Prichard, J. R. (2010), 'Sleep Patterns and Predictors of Disturbed Sleep in a Large Population of College Students', *Journal of Adolescent Health* **46**(2), 124–132.

Lundqvist, M., Carlsson, P., Sjödahl, R., Theodorsson, E. & Levin, L.-Å. (2017), 'Patient benefit of dog-assisted interventions in health care: a systematic review', *BMC Complement. Altern. Med.* **17**(1), 358.

Luque-Casado, A., Zabala, M., Morales, E., Mateo-March, M. & Sanabria, D. (2013), 'Cognitive Performance and Heart Rate Variability: The Influence of Fitness Level', *PLoS One* **8**(2), 1–9.

Lutwack-Bloom, P., Wijewickrama, R. & Smith, B. (2005), 'Effects of Pets versus People Visits with Nursing Home Residents', *Journal of Gerontological Social Work* **44**(3-4), 137–159.

Lyubomirsky, S., King, L. & Diener, E. (2005), 'The benefits of frequent positive affect: Does happiness lead to success?', *Psychological bulletin* **131**(6), 803.

Macaskill, A. (2013), 'The mental health of university students in the united kingdom', *British Journal of Guidance & Counselling* **41**(4), 426–441.

Mackersie, C. L. & Cones, H. (2011), 'Subjective and psychophysiological indexes of listening effort in a competing-talker task', *Journal of the American Academy of Audiology* **22**(2), 113–122.

Mah, L., Szabuniewicz, C. & Fiocco, A. J. (2016), 'Can anxiety damage the brain?', *Curr Opin Psychiatry* **29**(1), 56–63.

Mahmood, R., Parveen, N., Jillani, G., Safi, A. J., ud Din, S., ul Haq, I., ur Rehman, J. & ul Haq, A. (2006), 'Effect of Noise on Heart Rate', *Journal of Postgraduate Medicine* **20**(1), 12–15.

Mahmood, S. & Khatoon, T. (2011), 'Development and Validation of the Mathematics Anxiety Scale for Secondary and Senior Secondary School Students', *British Journal of Arts and Social Sciences* **2**(2), 169–179.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J. & Schwartz, P. J. (1996), 'Heart rate variability: Standards of measurement, physiological interpretation, and clinical use', *European Heart Journal* **17**(3), 354–381.

Mandryk, R. L. & Inkpen, K. M. (2004), 'Physiological indicators for the evaluation of co-located collaborative play', *Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04* pp. 1–10.

Marti, P., Pollini, A., Rullo, A. & Shibata, T. (2005), *in* 'Proceedings of the 2005 Annual Conference on European Association of Cognitive Ergonomics', University of Athens.

Matthews, G., Jones, D. M. & Chamberlain, A. G. (1990), 'Refining the measurement of mood: The UWIST Mood Adjective Checklist', *British Journal of Psychology* **81**(1), 17–42.

McCanne, T. R. & Hagstro m, A. H. (1996), 'Physiological hyperreactivity to stressors in physical child abusers and individuals at risk for being physically abusive', *Aggression and Violent Behavior* **1**(4), 345–358.

McGlynn, S. A., Geiskkovitch, D., Mitzner, T. L. & Rogers, W. A. (2016), PARO's stress-reduction potential for older adults, *in* 'Proceedings of the Human Factors and Ergonomics Society', pp. 1792–1796.

Melson, G. F., Kahn Jr, P. H., Beck, A. M., Friedman, B., Roberts, T. & Garrett, E. (2005), 'Robots as dogs? Children's interactions with the robotic dog AIBO and a live australian shepherd', *Computer-Human Interaction (CHI) Conference 2005* pp. 1649–1652.

Milicević, G. (2005), 'Low to high frequency ratio of heart rate variability spectra fails to describe sympatho-vagal balance in cardiac patients.', *Collegium antropologicum* **29**(1), 295–300.

Misra, R. & McKean, M. (2000), 'College Students' Academic Stress and Its Relation To Their Anxiety, Time Management, and Leisure Satisfaction', *American Journal of Health Studies* **16**, 11.

Mitsui, T., Shibata, T., Wada, K. & Tanie, K. (2001*a*), Psychophysiological effects by interaction with mental commit robot, *in* 'Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No.01CH37180)', Vol. 2.

Mitsui, T., Shibata, T., Wada, K. & Tanie, K. (2001*b*), Psychophysiological effects by interaction with mental commit robot, Vol. 2, IEEE.

Moerman, C. J., van der Heide, L. & Heerink, M. (2018), 'Social robots to support children's well-being under medical treatment: A systematic state-of-the-art review', *Journal of Child Health Care* p. 1367493518803031.

Mongillo, P., Pitteri, E., Adamelli, S., Bonichini, S., Farina, L. & Marinelli, L. (2015), 'Validation of a selection protocol of dogs involved in animal-assisted intervention', *Journal of Veterinary Behavior: Clinical Applications and Research* **10**(2), 103–110.

Mowbray, C. T., Megivern, D., Mandiberg, J. M., Strauss, S., Stein, C. H., Collins, K., Kopels, S., Curlin, C. & Lett, R. (2006), 'Campus mental health services: Recommendations for change', **76**(2), 226–237.

Moyle, W., Bramble, M., Jones, C. J., Murfield, J. E. & Bowers, B. J. (2017b), 'She Had a Smile on Her Face as Wide as the Great Australian Bite: A Qualitative Examination of Family Perceptions of a Therapeutic Robot and a Plush Toy', *The Gerontologist* .

Moyle, W., Jones, C., Cooke, M., O'Dwyer, S., Sung, B. & Drummond, S. (2013), 'Social robots helping people with dementia: Assessing efficacy of social robots in the nursing home environment', *2013 6th International Conference on Human System Interactions (HSI)* pp. 608–613.

Moyle, W., Jones, C. J., Murfield, J. E., Thalib, L., Beattie, E. R. A., Shum, D. K. H., O'Dwyer, S. T., Mervin, M. C. & Draper, B. M. (2017a), 'Use of a Robotic Seal as a Therapeutic Tool to Improve Dementia Symptoms: A Cluster-Randomized Controlled Trial', *Journal of the American Medical Directors Association* **18**(9), 766–773.

Moyle, W., Jones, C., Sung, B., Bramble, M., O'Dwyer, S., Blumenstein, M. & Estivill-Castro, V. (2016), 'What Effect Does an Animal Robot Called CuDDler Have on the Engagement and Emotional Response of Older People with Dementia? A Pilot Feasibility Study', *International Journal of Social Robotics* **8**(1), 145–156.

Münzel, T., Gori, T., Babisch, W. & Basner, M. (2014a), 'Cardiovascular effects of environmental noise exposure', *European heart journal* **35**(13), 829–836.

Münzel, T., Schmidt, F. P., Steven, S., Herzog, J., Daiber, A. & Sørensen, M. (2018b), 'Environmental noise and the cardiovascular system', *Journal of the American College of Cardiology* **71**(6), 688–697.

Myrick, J. G. (2015), 'Emotion regulation, procrastination, and watching cat videos online: Who watches internet cats, why, and to what effect?', *Computers in human behavior* **52**, 168–176.

Nagasawa, M., Kikusui, T., Onaka, T. & Ohta, M. (2009), 'Dog's gaze at its owner increases owner's urinary oxytocin during social interaction', *Hormones and Behavior* **55**(3), 434–441.

Nahm, N., Lubin, J., Lubin, J., Bankwitz, B. K., Castelaz, M., Chen, X., Shackson, J. C., Aggarwal, M. N. & Totten, V. Y. (2012), 'Therapy dogs in the emergency department', *Western journal of emergency medicine* **13**(4), 363.

Nasiri, P., Monazzam, M., Mohammadian, F., Zakerian, S. & Azam, K. (2014), 'Assessment of the effects of different noise levels on mental performance of individual doing mathematical calculation', *International Research Journal of Applied and Basic Sciences* **8**(7), 933–939.

Nguyen, T. A. & Zeng, Y. (2014), 'A physiological study of relationship between designer's mental effort and mental stress during conceptual design', *Computer-Aided Design* **54**, 3–18.

Oatley, K. & Jenkins, J. M. (1992), ‘Human Emotions: Function and Dysfunction’, *Annual Review of Psychology* **43**(1), 55–85.

O’Haire, M. (2010), ‘Companion animals and human health: benefits, challenges, and the road ahead’, *Journal of Veterinary Behavior: Clinical Applications and Research* **5**(5), 226–234.

O’haire, M. E., McKenzie, S. J., Beck, A. M. & Slaughter, V. (2015), ‘Animals may act as social buffers: Skin conductance arousal in children with autism spectrum disorder in a social context’, *Developmental psychobiology* **57**(5), 584–595.

Öhman, A. (2002), ‘Automaticity and the amygdala: Nonconscious responses to emotional faces’, *Current Directions in Psychological Science* **11**(2), 62–66.

Okita, S. Y. (2013), ‘Self-other’s perspective taking: the use of therapeutic robot companions as social agents for reducing pain and anxiety in pediatric patients.’, *Cyberpsychology, behavior and social networking* **16**(6), 1–6.

Osborne, J. W. & Overbay, A. (2004), ‘The power of outliers (and why researchers should always check for them)’, *Practical Assessment, Research & Evaluation* **9**(6), 1 – 8.

Out, D., Pieper, S., Bakermans-Kranenburg, M. & Van IJzendoorn, M. v. (2010), ‘Physiological reactivity to infant crying: a behavioral genetic study’, *Genes, Brain and Behavior* **9**(8), 868–876.

Oxford English Dictionary (2017), ‘Oxford English Dictionary Online’.
URL: http://dictionary.oed.com

Palermo, R. & Coltheart, M. (2004), ‘Photographs of facial expression: Accuracy, response times, and ratings of intensity’, *Behavior Research Methods, Instruments, and Computers* **36**(4), 634–638.

Palestrini, C., Calcaterra, V., Cannas, S., Talamonti, Z., Papotti, F., Buttram, D. & Pelizzo, G. (2017), ‘Stress level evaluation in a dog during animal-assisted therapy in pediatric surgery’, *Journal of Veterinary Behavior: Clinical Applications and Research* **17**, 221–222.

Pandey, S. & Bright, C. L. (2008), ‘What are degrees of freedom?’, *Social Work Research* **32**(2), 119–128.

Pannucci, C. J. & Wilkins, E. G. (2010), ‘Identifying and avoiding bias in research’, *Plastic and reconstructive surgery* **126**(2), 619.

Park, D., Ramirez, G. & Beilock, S. L. (2014), ‘The role of expressive writing in math anxiety.’, *Journal of Experimental Psychology: Applied* **20**(2), 103.

Park, E., Cho, M. & Ki, C.-S. (2009), ‘Correct use of repeated measures analysis of variance’, *The Korean journal of laboratory medicine* **29**(1), 1–9.

Park, H., Oh, S., Noh, Y., Kim, J. Y. & Kim, J.-H. (2018a), ‘Heart rate variability as a marker of distress and recovery: The effect of brief supportive expressive group therapy with mindfulness in cancer patients’, *Integrative cancer therapies* pp. 825–831.

Park, S. H., Lee, P. J. & Jeong, J. H. (2018b), ‘Effects of noise sensitivity on psychophysiological responses to building noise’, *Building and Environment* **136**, 302–311.

Paro Robots US, I. (2014), 'How to take care of Paro?'.

Parsons, R. (2002), 'Environmental psychophysiology', *Handbook of environmental psychology* **172**.

Passchier-Vermeer, W. & Passchier, W. F. (2000), 'Noise exposure and public health.', *Environmental health perspectives* **108**(Suppl 1), 123.

Pedersen, P. L. (2011), Do Elders Dream of Electric Seals ? A SCOT analysis of mental commitment robot PARO in elderly care, PhD thesis.

Perham, N., Hodgetts, H., Banbury, S. et al. (2013), 'Mental arithmetic and non-speech office noise: An exploration of interference-by-content', *Noise and Health* **15**(62), 73.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., Hedden, S. L., Crosby, A. E., Visser, S. N., Schieve, L. a., Parks, S. E., Hall, J. E., Brody, D., Simile, C. M., Thompson, W. W., Baio, J., Avenevoli, S., Kogan, M. D. & Huang, L. N. (2013), 'Mental health surveillance among children–United States, 2005-2011.', *Morbidity and mortality weekly report. Surveillance summaries (Washington, D.C. : 2002)* **62 Suppl 2**(2), 1–35.

Petersen, S., Houston, S., Qin, H., Tague, C. & Studley, J. (2017), 'The Utilization of Robotic Pets in Dementia Care', *Journal of Alzheimer's Disease* **55**(2), 569–574.

Petry, N. M. (2002), 'A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers', *The Gerontologist* **42**(1), 92–99.

Pflanzer, R. (2005), *Experimental and Applied Physiology Laboratory Manual*, McGraw-Hill Companies,Incorporated.

Picard, R. & Wolf, G. (2015), 'Guest Editorial Sensor Informatics and Quantified Self', *IEEE Journal of Biomedical and Health Informatics* **19**(5), 1531–1531.

Piéri, J. F., Crowe, J. a., Hayes-Gill, B. R., Spencer, C. J., Bhogal, K. & James, D. K. (2001), 'Compact long-term recorder for the transabdominal foetal and maternal electrocardiogram.', *Medical & biological engineering & computing* **39**(1), 118–125.

Pipitpukdee, J., Phantachat, W., Science, N., Development, T., Phantachat, W., Electronic, N. & Technology, C. (2011), 'The study of the pet robot therapy in Thai autistic children', *Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology (i-CREATe '11). Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre* pp. 43–46.

Poscia, A., Stojanovic, J., La Milia, D. I., Duplaga, M., Grysztar, M., Moscato, U., Onder, G., Collamati, A., Ricciardi, W. & Magnavita, N. (2018), 'Interventions targeting loneliness and social isolation among the older people: an update systematic review', *Experimental gerontology* **102**, 133–144.

Posner, J., Russell, J. A. & Peterson, B. S. (2005), 'The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology', *Development and Psychopathology* **17**(3), 715–734.

Pour, P. A., Hussain, M. S., AlZoubi, O., D'Mello, S. & Calvo, R. A. (2010), 'The impact of system feedback on learners' affective and physiological states BT - 10th International Conference on Intelligent Tutoring Systems, ITS 2010, June 14, 2010 - June 18, 2010', **6094 LNCS**(PART 1), 264–273.

Preuß, D. & Legal, F. (2017), 'Living with the animals: animal or robotic companions for the elderly in smart homes?', *Journal of medical ethics* **43**(6), 407–410.

Pu, L., Moyle, W., Jones, C. & Todorovic, M. (2018), 'The effectiveness of social robots for older adults: A systematic review and meta-analysis of randomized controlled studies', *The Gerontologist* .

Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., Kemp, A. H., D.S., Q., A.J., G., T., O., I.B., H. & A.H., K. (2012), 'Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition', *International Journal of Psychophysiology* **86**(2), 168–172.

Rabbitt, S. M., Kazdin, A. E. & Scassellati, B. (2015), 'Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use', *Clinical Psychology Review* **35**(2015), 35–46.

Ramanathan, M., Bhavanani, A. B. & Trakroo, M. (2017), 'Effect of a 12-week yoga therapy program on mental health status in elderly women inmates of a hospice', *International journal of yoga* **10**(1), 24.

Randall, T. T. (2017), 'Meet paro: The robotic seal that is making us more human'. [Online; accessed September 20, 2017].
URL: http://www.3rdactmagazine.com/wellness/health-wellness/paro-the-robotic-seal

Razali, N. M., Wah, Y. B. et al. (2011), 'Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests', *Journal of statistical modeling and analytics* **2**(1), 21–33.

Recio, A., Linares, C., Banegas, J. R. & Díaz, J. (2016), 'The short-term association of road traffic noise with cardiovascular, respiratory, and diabetes-related mortality', *Environmental research* **150**, 383–390.

Remington, R. (2002a), Hand massage in the agitated elderly, *in* 'Massage Therapy', Elsevier, pp. 165–185.

Roberts, M. (2011), *The everything guide to stress management: Step-by-step advice for elimination stress and living a happy*, Routledge History of Philosophy, Avon, Massachusetts, Adams Media.

Robinson, H., Broadbent, E. & MacDonald, B. (2015), 'Group sessions with Paro in a nursing home: Structure, observations and interviews', *Australasian Journal on Ageing* **35**(2), 106–112.

Robinson, H., MacDonald, B. A., Kerse, N. & Broadbent, E. (2013a), 'Suitability of healthcare robots for a dementia unit and suggested improvements.', *Journal of the American Medical Directors Association* **14**(1), 34–40.

Robinson, H., Macdonald, B. & Broadbent, E. (2013c), 'Physiological effects of a companion robot on blood pressure of older people in residential care facility: A pilot study.', *Australasian journal on ageing* **34**(1), 27–32.

Robinson, H., MacDonald, B., Kerse, N. & Broadbent, E. (2013b), 'The Psychosocial Effects of a Companion Robot: A Randomized Controlled Trial', *Journal of the American Medical Directors Association* **14**(9), 661–667.

Robotham, D. (2008), 'Stress among higher education students: Towards a research agenda', *Higher Education* **56**, 735–746.

Rodrigues, S., Paiva, J. S., Dias, D., Aleixo, M., Filipe, R. M. & Cunha, J. P. S. (2018), 'Cognitive impact and psychophysiological effects of stress using a biomonitoring platform', *International journal of environmental research and public health* **15**(6), 1080.

Rose, A. J. & Rudolph, K. D. (2006), 'A Review of Sex Differences in Peer Relationship Processes: Potential Trade-offs for the Emotional and Behavioral Development of Girls and Boys.', *Psychological Bulletin* **132**(1), 98–131.

Rothman, K. J. (1990), 'No adjustments are needed for multiple comparisons', *Epidemiology* pp. 43–46.

Russell, J. A. (1980), 'A circumplex model of affect', *Journal of Personality and Social Psychology* **39**(6), 1161–1178.

Rylander, R. (2006), 'Noise, stress and annoyance', *Noise & Vibration Worldwide* (June), 1–13.

Sabanovic, S., Bennett, C. C., Chang, W.-L. & Huber, L. (2013), 'Paro robot affects diverse interaction modalities in group sensory therapy for older adults with dementia', *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)* pp. 1–6.

Sahoo, R. & Sethi, S. (2015), 'Functional analysis of mental stress based on physiological data of GSR sensor', **337**, 109–117.

Saito, T., Shibata, T., Wada, K. & Tanie, K. (2003), Relationship between interaction with the mental commit robot and change of stress reaction of the elderly, *in* 'Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium', Vol. 1, pp. 119–124.

Sakakibara, M., Takeuchi, S. & Hayano, J. (1994), 'Effect of relaxation training on cardiac parasympathetic tone', *Psychophysiology* **31**(3), 223–228.

Saleem, S., Hussain, M. M., Majeed, S. M. I. & Khan, M. A. (2012), 'Gender differences of heart rate variability in healthy volunteers.', *JPMA. The Journal Of The Pakistan Medical Association* **62**(5), 422–425.

Sammito, S., Thielmann, B., Seibt, R., Klussmann, A., Weippert, M. & Böckelmann, I. (2015), 'Guideline for the application of heart rate and heart rate variability in occupational medicine and occupational science', pp. 1–29.

Santana, M. D. R., Martiniano, E. C., Monteiro, L. R. L., Valenti, V. E., Garner, D. M., Sorpreso, I. C. E. & Abreu, L. C. d. (2017), 'Musical auditory stimulation influences heart rate autonomic responses to endodontic treatment', *Evidence-Based Complementary and Alternative Medicine* **2017**, 1–7.

Saunders, J., Parast, L., Babey, S. H. & Miles, J. V. (2017), 'Exploring the differences between pet and non-pet owners: Implications for human-animal interaction research and policy', *PLoS One* **12**(6), 1–15.

Sawir, E., Marginson, S., Deumert, A., Nyland, C. & Ramia, G. (2008), 'Loneliness and International Students: An Australian Study', *Journal of Studies in International Education* **12**(2), 148–180.

Schiffrin, H. H. & Nelson, S. K. (2010), 'Stressed and happy? investigating the relationship between happiness and perceived stress', *Journal of Happiness Studies* **11**(1), 33–39.

Schwartz, S. (2003), 'Separation anxiety syndrome in dogs and cats', *Journal of the American Veterinary Medical Association* **222**(11), 1526–1532.

Seaward, B. (2012), 'Physiology of Estres', *Managing Stress* pp. 34–48.

Seifert, G., Kanitz, J.-L., Rihs, C., Krause, I., Witt, K. & Voss, A. (2018), 'Rhythmical massage improves autonomic nervous system function: a single-blind randomised controlled trial', *Journal of integrative medicine* **16**(3), 172–177.

Selvaraj, N., Jaryal, A., Santhosh, J., Deepak, K. K. & Anand, S. (2008), 'Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography', *Journal of Medical Engineering and Technology* **32**(6), 479–484.

Serpell, J. A., Coppinger, R., Fine, A. H. & Peralta, J. M. (2010), Welfare Considerations in Therapy and Assistance Animals, *in* 'Handbook on Animal-Assisted Therapy', pp. 481–503.

Shafer, D. & Zhang, Z. (2012), 'Beginning statistics v. 1.0', *O. New York, NY: Environmental Protection Agency* .

Shaffer, F. & Ginsberg, J. (2017), 'An overview of heart rate variability metrics and norms', *Frontiers in public health* **5**, 258.

Shankar, S. & Singh, R. (2014), 'Demystifying statistics how to choose a statistical test?', *Indian Journal of Rheumatology* **9**(2), 77–81.

Shearer, A., Hunt, M., Chowdhury, M. & Nicol, L. (2016), 'Effects of a brief mindfulness meditation intervention on student stress and heart rate variability.', *International Journal of Stress Management* **23**(2), 232.

Sherman, G. D., Haidt, J. & Coan, J. A. (2009), 'Viewing Cute Images Increases Behavioral Carefulness', *Emotion* **9**(2), 282–286.

Shi, H., Yang, L., Zhao, L., Su, Z., Mao, X., Zhang, L. & Liu, C. (2017), 'Differences of heart rate variability between happiness and sadness emotion states: a pilot study', *Journal of Medical and Biological Engineering* **37**(4), 527–539.

Shibata, T. (2012), 'Therapeutic seal robot as biofeedback medical device: Qualitative and quantitative evaluations of robot therapy in dementia care', *Proceedings of the IEEE* **100**(8), 2527–2538.

Shibata, T. & Coughlin, J. F. (2014), 'Trends of robot therapy with neurological therapeutic seal robot, PARO', *Journal of Robotics and Mechatronics* **26**(4), 418–425.

Shibata, T., Kawaguchi, Y. & Wada, K. (2011), 'Investigation on People Living with Seal Robot at Home', *International Journal of Social Robotics* **4**(1), 53–63.

Shibata, T., Wada, K., Ikeda, Y. & Sabanovic, S. (2009), 'Cross-Cultural Studies on Subjective Evaluation of a Seal Robot', *Advanced Robotics* **23**(4), 443–458.

Shibata, T., Wada, K. & Tanie, K. (2003), 'Statistical analysis and comparison of questionnaire results of subjective evaluations of seal robot in Japan and U.K.', *Proceedings - IEEE International Conference on Robotics and Automation* **3**, 52–57.

Shields, M. (2004), 'Stress, health anfd the benefit of social support', *Health Reports* **15**(1), 9–37.

Shiloh, S., Sorek, G. & Terkel, J. (2003), 'Reduction of state-anxiety by petting animals in a controlled laboratory experiment', *Anxiety, Stress and Coping* **16**(4), 387–395.

Shin, J.-H., Lee, J.-Y., Yang, S.-H., Lee, M.-Y. & Chung, I.-S. (2016), 'Factors related to heart rate variability among firefighters', *Annals of occupational and environmental medicine* **28**(1), 25.

Silva, L. C. D. E. & Miyasato, I. T. (1997), Facial emotion recognition using multi-modal information, *in* 'Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat.', Vol. 1, IEEE, pp. 397–401.

Silvia, P. J. & Warburton, J. B. (2006), 'Positive and negative affect: Bridging states and traits', *Comprehensive handbook of personality and psychopathology* **1**, 268–284.

Sim, C. S., Sung, J. H., Cheon, S. H., Lee, J. M., Lee, J. W. & Lee, J. (2015), 'The Effects of Different Noise Types on Heart Rate Variability in Men', *Yonsei Medical Journal* **56**(1), 235.

Sin, N. L. & Lyubomirsky, S. (2009), 'Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis', *Journal of clinical psychology* **65**(5), 467–487.

Smallwood, J., O'Connor, R. C., Sudbery, M. V. & Obonsawin, M. (2007), 'Mind-wandering and dysphoria', *Cognition and Emotion* **21**(4), 816–842.

Smith, A., Waters, B., Jones, H. et al. (2010), 'Effects of prior exposure to office noise and music on aspects of working memory', *Noise and Health* **12**(49), 235.

Somnox (2017), 'Somnox sleep robotscientific background', pp. 1–21.
URL: http://somnox.nl/files/science/SomnoxScientificValidation_v4.pdf

Sorrell, J. M., Cangelosi, P. R. & Sorrell, J. M. (2010), 'Walking for Therapy with Man's Best Friend', *Journal of Psychosocial Nursing and Mental Health Services* **48**(3), 19–22.

Spielberger, C. D. (1987), ‘State-Trait Anxiety Inventory’.

Spielberger, C. & Sarason, I. (2013), *Stress And Emotion: Anxiety, Anger, & Curiosity*, Routledge History of Philosophy, Taylor & Francis.

Sripongngam, T., Eungpinichpong, W., Sirivongs, D., Kanpittaya, J., Tangvoraphonkchai, K. & Chanaboon, S. (2015), ‘Immediate effects of traditional thai massage on psychological stress as indicated by salivary alpha-amylase levels in healthy persons’, *Medical science monitor basic research* **21**, 216.

Stansfeld, S. A. & Matheson, M. P. (2003), ‘Noise pollution: Non-auditory effects on health’, *British Medical Bulletin* **68**, 243–257.

Stubbs, B., Veronese, N., Vancampfort, D., Prina, A. M., Lin, P.-Y., Tseng, P.-T., Evangelou, E., Solmi, M., Kohler, C., Carvalho, A. F. et al. (2017), ‘Perceived stress and smoking across 41 countries: A global perspective across europe, africa, asia and the americas’, *Scientific reports* **7**(1), 7597.

Sullivan, G. M. & Feinn, R. (2012), ‘Using effect size—or why the p value is not enough’, *Journal of graduate medical education* **4**(3), 279–282.

Sung, H.-C., Chang, S.-M., Chin, M.-Y. & Lee, W.-L. (2014), ‘Robot-assisted therapy for improving social interactions and activity participation among institutionalized older adults: A pilot study.’, *Asia-Pacific psychiatry : official journal of the Pacific Rim College of Psychiatrists* pp. 1–6.

Sung, J. H., Lee, J., Jeong, K. S., Lee, S., Lee, C., Jo, M.-W. & Sim, C. S. (2017), ‘Influence of transportation noise and noise sensitivity on annoyance: a cross-sectional study in south korea’, *International journal of environmental research and public health* **14**(3), 322.

Szabo, S., Tache, Y. & Somogyi, A. (2012), ‘The legacy of Hans Selye and the origins of stress research: A retrospective 75 years after his landmark brief “Letter” to the Editor # of *Nature*’, *Stress* **15**(5), 472–478.

Szalma, J. & Hancock, P. (2010), A meta-analytic review of the effects of noise on performance: Moderating effects of task and noise characteristics, *in* ‘Proceedings of the Human Factors and Ergonomics Society Annual Meeting’, Vol. 54, SAGE Publications Sage CA: Los Angeles, CA, pp. 1660–1664.

Szalma, J. L. & Hancock, P. A. (2011), ‘Noise effects on human performance: a meta-analytic synthesis.’, *Psychological bulletin* **137**(4), 682.

Taelman, J., Vandeput, S., Spaepen, a. & Huffel, S. V. (2009), ‘Influence of Mental Stress on Heart Rate and Heart Rate Variability’, *Ecifmbe 2008* **29**(1), 1366–1369.

Taggart, W., Turkle, S. & Kidd, C. D. (2005), ‘An Interactive Robot in a Nursing Home : Preliminary Remarks Field Setting : Nursing Homes’, *Towards Social Mechanisms of Android Science: A COGSCI Workshop* .

Takayanagi, K., Kirita, T. T. & Shibata, T. (2014), ‘Comparison of Verbal and Emotional Responses of Elderly People with Mild/Moderate Dementia and Those with Severe Dementia in Responses to Seal Robot, PARO.’, *Frontiers in aging neuroscience* **6**(September), 257.

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O. & Karjalainen, P. A. (2014), 'Kubios HRV - Heart rate variability analysis software', *Computer Methods and Programs in Biomedicine* **113**(1), 210–220.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996), 'Heart Rate Variability : Standards of Measurement, Physiological Interpretation, and Clinical Use', *Circulation* **93**(5), 1043–1065.

Tatchell, T., Jordan, T. R., Waite, P. J. & Tatchell, R. H. (2004), 'Transmissive reminiscence therapy with college students and institutionalized senior adults', *Journal of Intergenerational Relationships* **1**(4), 35–52.

Terry, P. C., Lane, a. M. & Fogarty, G. J. (2003), 'Construct validity of the Profile of Mood States - Adolescents for use with adults', *Psychology of Sport and Exercise* **4**(2), 125–139.

Thodberg, K., Sørensen, L. U., Christensen, J. W., Poulsen, P. H., Houbak, B., Damgaard, V., Keseler, I., Edwards, D. & Videbech, P. B. (2016), 'Therapeutic effects of dog visits in nursing homes for the elderly: Effects of dog visits in nursing homes', *Psychogeriatrics* **16**(5), 289–297.

Thoma, M. V., La Marca, R., Brönnimann, R., Finkel, L., Ehlert, U. & Nater, U. M. (2013), 'The effect of music on the human stress response', *PloS one* **8**(8), 1–12.

Tiberio, L., Cesta, A. & Belardinelli, M. (2013), 'Psychophysiological Methods to Evaluate User's Response in Human Robot Interaction: A Review and Feasibility Study', *Robotics* **2**(2), 92–121.

Trammell, J. P. (2017), 'The effect of therapy dogs on exam stress and memory', *Anthrozoös* **30**(4), 607–621.

Travis, D. (2002), *E-Commerce Usability: Tools and Techniques to Perfect the On-Line Experience (1st ed.)*, Routledge, New York.

Tsai, C. C., Friedmann, E. & Thomas, S. A. (2010), 'The effect of animal-assisted therapy on stress responses in hospitalized children', *Anthrozoos* **23**(3), 245–258.

Turkle, S., Taggart, W., Kidd, C. D. & Dasté, O. (2006), 'Relational artifacts with children and elders: the complexities of cybercompanionship', *Connection Science* **18**(4), 347–361.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A. & Zelson, M. (1991), 'Stress recovery during exposure to natural and urban environments', *Journal of environmental psychology* **11**(3), 201–230.

van Erp, J. B. F. & Toet, A. (2015), 'Social Touch in Human Computer Interaction', *Frontiers in Digital Humanities* **2**.

van Ravenswaaij-Arts, C. M. A. (1993), 'Heart Rate Variability', *Annals of Internal Medicine* **118**(6), 436.

Vanitha, L. & Suresh, G. R. (2014), Hierarchical SVM to detect mental stress in human beings using Heart Rate Variability, *in* '2014 2nd International Conference on Devices, Circuits and Systems (ICDCS)', pp. 1–5.

Venables, P. H. & Mitchell, D. a. (1996), 'The effects of age, sex and time of testing on skin conductance activity', *Biological Psychology* **43**(2), 87–101.

Vescio, B., Salsone, M., Gambardella, A. & Quattrone, A. (2018), 'Comparison between electrocardiographic and earlobe pulse photoplethysmographic detection for evaluating heart rate variability in healthy subjects in short- and long-term recordings', *Sensors* **18**(3).

von Baeyer, C. L., Piira, T., Chambers, C. T., Trapanotto, M. & Zeltzer, L. K. (2005), 'Guidelines for the cold pressor task as an experimental pain stimulus for use with children', *The journal of Pain* **6**(4), 218–227.

von Dawans, B., Fischbacher, U., Kirschbaum, C., Fehr, E. & Heinrichs, M. (2012), 'The social dimension of stress reactivity: acute stress increases prosocial behavior in humans.', *Psychol Sci* **23**(6), 651–660.

Vormbrock, J. K. & Grossberg, J. M. (1988), 'Cardiovascular effects of human-pet dog interactions', *Journal of Behavioral Medicine* **11**(5), 509–517.

Voss, A., Heitmann, A., Schroeder, R., Peters, A. & Perz, S. (2012), 'Short-term heart rate variability—age dependence in healthy subjects', *Physiological Measurement* **33**(8), 1289–1311.

Wada, K. & Shibata, T. (2007a), 'Robot therapy in a care house - Change of relationship among the residents and seal robot during a 2-month long study', *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication* (October), 107–112.

Wada, K. & Shibata, T. (2007b), Living with seal robots - Its sociopsychological and physiological influences on the elderly at a care house, *in* 'IEEE Transactions on Robotics', Vol. 23, pp. 972–980.

Wada, K., Shibata, T., Musha, T. & Kimura, S. (2005b), Effects of Robot Therapy for Demented Patients Evaluated by EEG, IEEE, pp. 1552–1557.

Wada, K., Shibata, T., Musha, T. & Kimura, S. (2008), 'Robot therapy for elders affected by dementia', *IEEE Engineering in Medicine and Biology Magazine* **27**(4), 53–60.

Wada, K., Shibata, T., Saito, T. & Tanie, K. (2004*a*), Effects of robot-assisted activity for elderly people and nurses at a day service center, *in* 'Proceedings of the IEEE', Vol. 92, pp. 1780–1788.

Wada, K., Shibata, T., Saito, T. & Tanie, K. (2004*b*), 'Effects of three months robot assisted activity to depression of elderly people who stay at a health service facility for the aged', *SICE 2004 Annual Conference* pp. 2709–2714.

Wada, K., Takasawa, Y. & Shibata, T. (2013), 'Robot therapy at facilities for the elderly in Kanagawa prefecture - A report on the experimental result of the first week', *IEEE RO-MAN: The 22nd IEEE International Symposium on Robot and Human Interactive Communication Gyeongju, Korea* pp. 757–761.

Wada, K. W. K., Shibata, T. S. T., Saito, T. S. T., Sakamoto, K. S. K. & Tanie, K. T. K. (2005a), 'Robot assisted activity at a health service facility for the aged for 17 months:

an interim report of long-term experiment', *IEEE Workshop on Advanced Robotics and its Social Impacts, 2005.* pp. 127–132.

Wagemaker, E., Dekkers, T. J., Agelink van Rentergem, J. A., Volkers, K. M. & Huizenga, H. M. (2017), 'Advances in mental health care: five n= 1 studies on the effects of the robot seal paro in adults with severe intellectual disabilities', *Journal of Mental Health Research in Intellectual Disabilities* **10**(4), 309–320.

Waghachavare, V. B., Dhumale, G. B., Kadam, Y. R. & Gore, A. D. (2013), 'A study of stress among students of professional colleges from an urban area in India', *Sultan Qaboos University Medical Journal* **13**(3), 422–429.

Walker, E. D., Brammer, A., Cherniack, M. G., Laden, F. & Cavallari, J. M. (2016), 'Cardiovascular and stress responses to short-term noise exposures—a panel study in healthy males', *Environmental research* **150**, 391–397.

Wang, H. M. & Huang, S. C. (2012), 'SDNN/RMSSD as a surrogate for LF/HF: A revised investigation', *Modelling and Simulation in Engineering* **2012**, 1–8.

Wang, J., Korczykowski, M., Rao, H., Fan, Y., Pluta, J., Gur, R. C., McEwen, B. S. & Detre, J. a. (2007), 'Gender difference in neural response to psychological stress', *Social Cognitive and Affective Neuroscience* **2**(3), 227–239.

Ward, R. D. & Marsden, P. H. (2003), 'Physiological responses to different web page designs', *Int. J. Hum.-Comput. Stud.* **59**(1-2), 199–212.

Watkins, J. C. (2016), 'An introduction to the science of statistics: From theory to implementation'.

Westeyn, T., Presti, P. & Starner, T. (2007), 'ActionGSR: A combination galvanic skin response-accelerometer for physiological measurements in active environments', *Proceedings - International Symposium on Wearable Computers, ISWC* (Figure 2), 129–130.

Westman, J. C. & Walters, J. R. (1981), 'Noise and stress: A comprehensive approach', *Environmental Health Perspectives* **Vol. 41**(October), 291–309.

Wilson, C. C. & Barker, S. B. (2003), 'Challenges in Designing Human-Animal Interaction Research', *American Behavioral Scientist* **47**(1), 16–28.

Wohlfarth, R. & Sandstedt, L. (2016), *Animal Assisted Activities with Dogs: Guideline for Basic Requirements & Knowledge*, Publishing House of Janusz Korczak Pedagogical University.

Wolff, A. I. & Frishman, W. H. (2004), 'Animal-assisted therapy in cardiovascular disease', *Seminars in Integrative Medicine* **2**(4), 131–134.

Wood, E., Ohlsen, S., Thompson, J., Hulin, J. & Knowles, L. (2018), 'The feasibility of brief dog-assisted therapy on university students stress levels: the paws study', *Journal of Mental Health* **27**(3), 263–268.

Wood, N., Sharkey, A., Mountain, G. & Millings, A. (2015), 'The Paro robot seal as a social mediator for healthy users', *4th International Symposium on New Frontiers in Human Robot Interaction* .

Wood, R., Maraj, B., Lee, C. M. & Reyes, R. (2002), 'Short-term heart rate variability during a cognitive challenge in young and older adults', *Age and ageing* **31**(2), 131–135.

World Health Organization (2010), 'People with mental disabilities cannot be forgotten'.
URL: http://www.who.int/mediacentre/news/releases/2010/mental_disabilities_20100916/en/

World Health Organization (2014), 'Mental health: strengthening our response'.
URL: http://www.who.int/mediacentre/factsheets/fs220/en/

Wormald, D., Lawrence, A. J., Carter, G. & Fisher, A. D. (2017), 'Reduced heart rate variability in pet dogs affected by anxiety-related behaviour problems', *Physiology and Behavior* **168**, 122–127.

Yamamoto, S. & Kimura, R. (2007), 'Investigation of playing with entertainment robotic pet of pre-school aged child in nursery school by video observation', *Proceedings of the SICE Annual Conference* pp. 654–659.

Yang, G., Baad-Hansen, L., Wang, K., Xie, Q. F. & Svensson, P. (2014), 'Effect of negative emotions evoked by light, noise and taste on trigeminal thermal sensitivity', *Journal of Headache and Pain* **1–12**(1).

Yang, J. (2015), 'Meet paro, a furry friend to dementia patients'. [Online; accessed January 21, 2016].
URL: http://onlinelibrary.wiley.com/doi/10.1002/9780470479216.corpsy0943/full

Yawn, B. P. (2008), 'Factors accounting for asthma variability: achieving optimal symptom control for individual patients', *Primary Care Respiratory Journal* **17**(3), 138.

Yu, R., Hui, E., Lee, J., Poon, D., Ng, A., Sit, K., Ip, K., Yeung, F., Wong, M., Shibata, T. & Woo, J. (2015), 'Use of a Therapeutic, Socially Assistive Pet Robot (PARO) in Improving Mood and Stimulating Social Interaction and Communication for People With Dementia: Study Protocol for a Randomized Controlled Trial.', *JMIR research protocols* **4**(2), 1–12.

Zamir, T. (2006), 'The moral basis of animal-assisted therapy', *Society and Animals* **14**(2), 179–199.

Appendix A

Appendix

A.1 Advertisement for study 1

[student-volunteers] Can interacting with a robot reduce your stress? Inbox x

Department of Computer Science raminuddin1@sheffield.ac.uk via lists.shef.ac.uk 11/04/2016
to student-volunt.

Dear Volunteer,

I would like to invite you to take part in my study, and have a chance to win a #50 gift card. This study attempts to understand how interactions with Paro, a robotic seal will affect you.
The study will take place in the Department of Computer Science Research Laboratory (Map: https://goo.gl/r0Fm0v) at the University of Sheffield and will last for 20 minutes.
During that time you'll be asked to listen to a sound and interact with the Paro robot whilst being connected to physiological sensors that record and measure your physiological reactions. The sensors are completely non-invasive and you'll be wearing the sensors on your fingertips on your non-dominant hand.

However, in order to ensure consistent results, if you have hearing difficulties; heart and skin health conditions; are taking prescribed medication; or have mental health problems then you should not participate in this study. This study is the same as the one previously advertised on 2nd March, so please do not sign up if you have already taken part.

If you'd like to take part in this study, please take a few short minutes to fill out the survey form with general information about yourself by clicking on this link: https://goo.gl/WP1NiN. At the end of the survey, you will be redirected to a booking page where you can book in for a session at a date and time that suits you.

This study is being conducted by Raihah Aminuddin. If you have any questions about this study, please email me at raminuddin1@sheffield.ac.uk.

All responses will be kept strictly anonymous and confidential. You may withdraw from the study at any point before or during your participation. This study has been approved by the Department of Computer Science Research Ethics Committee and is being carried out under the supervision of Amanda Sharkey (a.sharkey@sheffield.ac.uk).

Information related to this message is available at https://goo.gl/WP1NiN.

For information about this email list, including how to remove your name, please visit https://www.sheffield.ac.uk/cics/email/distributionlists.html and click the list name.

Advertisement for study 1: Email sent by Student Services information Desk (SSiD).

A.2 Advertisement for study 2

[student-volunteers] Does interacting with a Paro seal robot reduce your stress?

Inbox x

Rai Aminuddin raminuddin1@sheffield.ac.uk via lists.shef.ac.uk 6 Feb

to student-volunt.

If you are aged between 18 and 25, I would like to invite you to take part in my study, and have a chance to win a #15 gift card. If you are a Psychology student participating for course credits, you will receive 2 credits.

This study attempts to understand how interactions with Paro, a robotic seal will affect you.

The study will take place in the Department of Psychology (362 Mushroom Ln, Sheffield S10 2TS) (Map: http://tinyurl.com/z9klquv ; Picture: http://tinyurl.com/hpm3rwt) at the University of Sheffield and will last for 30 minutes.

During that time, you'll be asked to interact with the Paro robot and listen to an aversive sound whilst being connected to physiological sensors that record and measure your physiological reactions. The sensors are completely non-invasive and you'll be wearing the sensors on your fingertips on your non-dominant hand.

However, in order to ensure consistent results, if you have hearing difficulties, heart and skin health conditions, are taking prescribed medication, or have mental health problems then you should not participate in this study.

If you'd like to take part in this study, please use this link: http://tinyurl.com/gvfo8ch, you will be redirected to a booking page where you can book in for a session at a date and time that suits you.

If you are a Psychology student participating for course credits, please use Online Research Participation System (http://tinyurl.com/zojuxuu) to book your session.

This study is being conducted by Raihah Aminuddin. If you have any questions about this study, please email me at raminuddin1@sheffield.ac.uk.

All responses will be kept strictly anonymous and confidential. You may withdraw from the study at any point before or during your participation. This study has been approved by the Department of Computer Science Research Ethics Committee and is being carried out under the supervision of Amanda Sharkey (a.sharkey@sheffield.ac.uk).

For information about this email list, including how to remove your name, please visit https://www.sheffield.ac.uk/cics/email/distributionlists.html and click the list name.

Advertisement for study 2: Email sent by Student Services information Desk (SSID).

A.3 Advertisement for study 3

[student-volunteers] Feeling stressed? Would you like to play with a Paro robot? Receive 5GBP for taking part in a study Inbox x

Department of Computer Science raminuddin1@sheffield.ac.uk via lists.shef 9 May

to student-volunt.

If you are aged between 18 and 35, I would like to invite you to take part in my study. You will receive #5 as a reward for your participation. This experiment will take 30 - 45 minutes to complete.

This study attempts to understand how interactions with Paro, a robotic seal, will affect you.

During the experiment, you will be asked to complete a short questionnaire that is related to your current feelings, and your responses will be recorded using physiological sensors. The sensors will be attached on your fingers, chest and waist (but don't worry, you don't have to remove any clothes or headscarf).

There will also be video recorded during the robot session, but only if you give your permission.

You also will be asked to answer some simple mathematical questions while listening to an aversive sound delivered through headphones. You will be provided with paper and pencil to write down the calculation.

In order to ensure consistent results, if you have hearing difficulties, heart and skin health conditions, are taking prescribed medication, or have mental health problems then you should not participate in this study.

PLEASE NOTE: This study is the same as the one previously advertised on 5th April,
PLEASE DO NOT SIGN UP IF YOU HAVE ALREADY TAKEN PART.

If you'd like to take part in this study, please take a few short minutes to fill out the survey form with general information about yourself by clicking on this link: https://tinyurl.com/n76snxo. At the end of the survey, you will be redirected to a booking page where you can book in for a session at a date and time that suits you.

The study will take place in the Department of Computer Science (211 Portobello, Sheffield S1 4DP) (Map: https://tinyurl.com/lyna35g , Picture: https://tinyurl.com/lxcwc2x).

This study is being conducted by Raihah Aminuddin. If you have any questions about this study, please email me at raminuddin1@sheffield.ac.uk.

All responses will be kept strictly anonymous and confidential. You may withdraw from the study at any point before or during your participation. This study has been approved by the Department of Computer Science Research Ethics Committee and is being carried out under the supervision of Amanda Sharkey (a.sharkey@sheffield.ac.uk).

For information about this email list, including how to remove your name, please visit https://www.sheffield.ac.uk/cics/email/distributionlists and click the list name.

Advertisement for study 3: Email sent by Student Services information Desk (SSiD).

A.4 Participants Information Sheet for study 1

PARTICIPANT INFORMATION SHEET FOR STRESS AND ROBOT STUDY

INTRODUCTION

You are invited to take part in a research project. Before you decide whether to take part or not, it is important that you understand why the research is being conducted and what it will involve. Please take time to read the following information carefully. Feel free to ask me (raminuddin1@shef.ac.uk) if you would like clarification or more information about this study. Please take time to decide whether or not you wish to take part in the experiment. Thank you for reading this.

PURPOSE OF THIS STUDY

In this study, Paro, a robotic seal will be used. Paro was developed purposely to provide comfort in elderly people with Dementia. The aim of this study is to gain a better understanding of the reasons for Paro's effects on people. This study will help us to evaluate which aspects of the Paro robot are responsible for any effects that it may have.

However, if you have hearing difficulties; heart and skin health conditions; are taking prescribed medication; or have mental health problems then you should not participate in this study.

EXPERIMENTAL PROCEDURES

The study will be conducted in the Listening booth (Research Laboratory 3) in Department of Computer Science, University of Sheffield and will last for 30 minutes.

All instructions will be displayed on the computer screen. During the experiment, you will be asked to complete a short questionnaire that is related to your current feelings; and your responses will be recorded using skin conductance. The skin conductance sensors will be attached onto your fingertips.

Before the experiment begins, the researcher will explain what will happen during the sessions. The researcher will also show you how to interact with the robot. You also will be asked to listen to an aversive sound delivered through headphones that will be connected to a laptop. However, if you find the sound is too loud, you can take off the headphones and ask the experimenter to stop the experiment immediately.

The study consists of two sessions:
- First session, you will be asked to listen to an aversive sound for 3 minutes.
- Second session, you will be asked to interact with a robot for 3 minutes.

At the end of the experiment, you will be given the opportunity to ask further questions related to this study design and aim of the study.

Participants Information Sheet for study 1 (Page 1).

QUESTIONS OR COMPLAINTS
If you have any further questions or complaints, please contact:
Raihah Aminuddin (raminuddin1@shef.ac.uk)
or Amanda Sharkey (a.sharkey@shef.ac.uk).
The Department of Computer Science, Regent Court, 211 Portobello, Sheffield, S1 4DP.

If you feel your complaint has not been handled to your satisfaction, please contact the University of Sheffield's Registrar and Secretary at: (Office of the Registrar and Secretary, Firth Court, Western Bank, Sheffield, S10 2TN).

DATA CONFIDENTIALITY AND STORAGE
Your answers from the questionnaire and recorded media will be stored safely with the data collected through the experiment. Your personal information will be stored separately and not associated with the questionnaire/experiment data at any time. You will not be able to be identified in any reports or publications.

ETHICAL APPROVAL
This project has been ethically approved via the Department of Computer Science ethics review procedure.

RECORDED MEDIA
The audio and/or video recordings of your activities made during this research will be used only for analysis. They could be used for illustration in conference presentations, but if only if you agree to this use. No other use will be made of them without your written permission, and no one outside the project will be allowed access to the original recordings.

RECORDED BIOSIGNAL MEASUREMENT
In this study, skin conductance measurement will be used to collect your physiological responses. The procedure will involve you sitting down and sensors will be attached directly to your fingertips in a material similar to a foam-like plaster. This will be wired up to the system which will be wired up to a computer. There may be some slight discomfort with the removal of the sensors from your skin because the sensors are just very sticky.

Once again, thank you for taking the time to read this information sheet and for considering taking part in this study.

Participants Information Sheet for study 1 (Page 2).

A.5 Participants Information Sheet for study 2

PARTICIPANT INFORMATION SHEET FOR STRESS AND ROBOT STUDY

INTRODUCTION

You are invited to take part in a research project. Before you decide whether to take part or not, it is important that you understand why the research is being conducted and what it will involve. Please take time to read the following information carefully. Feel free to ask me (raminuddin1@shef.ac.uk) if you would like clarification or more information about this study. Please take time to decide whether or not you wish to take part in the experiment. Thank you for reading this.

PURPOSE OF THIS STUDY

In this study, Paro, a robotic seal will be used. Paro was developed purposely to provide comfort for elderly people with Dementia.

The aim of this study is to gain a better understanding of the reasons for Paro's effects on people. This study will help us to investigate any effect the Paro robot may have on stress.

However, if you have hearing difficulties, heart and skin health conditions, are taking prescribed medication, or have mental health problems then you should not participate in this study.

EXPERIMENTAL PROCEDURES

The study will be conducted in Department of Psychology, University of Sheffield and will last for 30 minutes.

All instructions will be displayed on the computer screen. During the experiment, you will be asked to complete a short questionnaire that is related to your current feelings, and your responses will be recorded using physiological sensors. The physiological sensors will be attached onto your fingertips.

Before the experiment begins, the researcher will explain what will happen during the sessions. The researcher will also show or tell you how to interact with the robot. You also will be asked to listen to an aversive sound delivered through headphones. However, if you find the sound is too loud, you can take off the headphones and ask the experimenter to stop the experiment immediately.

The study consists of two sessions:

(i) Robot session (5 minutes)

(ii) An aversive sound session (5 minutes).

At the end of the experiment, you will be given the opportunity to ask further questions related to this study design and aim of the study.

Participants Information Sheet for study 2 (Page 1).

A.6 Participants Information Sheet for study 3

PARTICIPANT INFORMATION SHEET FOR STRESS AND ROBOT STUDY

INTRODUCTION

You are invited to take part in a research project. Before you decide whether to take part or not, it is important that you understand why the research is being conducted and what it will involve. Please take time to read the following information carefully. Feel free to ask me (raminuddin1@shef.ac.uk) if you would like clarification or more information about this study. Please take time to decide whether or not you wish to take part in the experiment. Thank you for reading this.

PURPOSE OF THIS STUDY

In this study, Paro, a robotic seal will be used. Paro was developed purposely to provide comfort for elderly people with Dementia.

The aim of this study is to gain a better understanding of the reasons for Paro's effects on people. This study will help us to investigate any effect the Paro robot may have on stress.

However, if you have hearing difficulties, heart and skin health conditions, are taking prescribed medication, or have mental health problems then you should not participate in this study.

EXPERIMENTAL PROCEDURES

The study will be conducted in Listening booth (Research lab 3) in Department of Computer Science at University of Sheffield and will last for **30 - 45 minutes**.

All instructions will be displayed on the computer screen. During the experiment, you will be asked to complete a few short questionnaires about your current feelings, and your responses will be recorded using physiological sensors. **The physiological sensors will be attached to different parts of your body (e.g. fingers, chest and waist). You also will be asked to wear a belt around your waist. The belt will be placed over your clothes.**

Before the experiment begins, the researcher will explain what will happen during the sessions. The researcher will tell you how to interact with the robot. You also will be asked to answer 9 GSCE level mathematical questions (involving multiplication, division, addition, subtraction) while listening to an aversive sound delivered through headphones. However, if you find the sound is too loud, you can take off the headphones and ask the experimenter to stop the experiment immediately. You also will be provided papers and a pencil to write down the calculation.

The study consists of two sessions: **(i) Robot session (5 minutes)** and **(ii) task session (5 minutes)**.

At the end of the experiment, you will be given the opportunity to ask further questions related to this study design and aim of the study.

Participants Information Sheet for study 3 (Page 1).

A.7 Consent form

Consent form for stress and robot study

Name of researcher: Raihah Aminuddin (raminuddin1@shef.ac.uk)

Participant identification number (to be filled in by researcher)	

Please answer the following questions:

	Yes	No
I confirm that I have read and understand the information sheet for this study. I have had the opportunity to consider the information and ask questions.		
My participation is voluntary and I am free to withdraw at any time without giving any reason, without consequence.		
All my collected data will be treated as confidential. I agree to my anonymous data being used in future research.		
If I have any questions regarding the study, I can contact the researcher at raminuddin1@sheffield.ac.uk, and her supervisor at a.sharkey@shef.ac.uk.		
I agree to take part in this study.		

Video Recording

As part of this study to assist with accurate recording of your responses, a video recording will be made during your participation in this study.

You have the right to refuse and stop the video recording at any time and the entire tape or any portion of the tape may be erased at your request.

	Yes	No
I understand that video recording will take place in this study.		
I understand that the recording(s) will be stored securely in appropriate file formats on servers belonging to the University.		

Please indicate below the uses of these video recording to which you are willing to consent. In any use of these video recording, your name will not be identified.

	Yes	No
I consent to the video recording being studied by the investigator.		
I consent to the photograph/video recording being used in public or scientific presentations. I also understand that my face will not be visible in the photo/video during the presentations.		

Participant's details:

Name of participant	
University email address	
Signature (Please put your initials)	
Date	4 Sep 2017

Consent form.

A.8 Positive and Negative Affect Questionnaire

This scale consists of a number of words that describe different feelings and emotions. Read each item and indicate to what extent you feel this way right now, that is, at the present moment.

	1	2	3	4	5
	Very slightly or Not at All	A Little	Moderately	Quite a Bit	Extremely
Interested					
Distressed					
Excited					
Upset					
Strong					
Guilty					
Scared					
	Very slightly or Not at All	A Little	Moderately	Quite a Bit	Extremely
Hostile					
Enthusiastic					
Proud					
Irritable					
Alert					
Ashamed					
Inspired					
	Very slightly or Not at All	A Little	Moderately	Quite a Bit	Extremely
Nervous					
Determined					
Attentive					
Jittery					
Active					
Afraid					

Positive and Negative Affect Questionnaire.

A.9 Neutral reading material

Chapter 1

Musical Values

A Television Commercial

'I want to be . . . *a musician*.' Those are the opening words in a television commercial for Prudential pension plans which was being broadcast in late 1992. It begins with a young man sitting back in a chair, a dreamy, wistful expression on his face, listening to music on headphones (Fig. 1). He is absorbed in the music; he taps his foot and bobs his head in time to it. And yet he is not completely taken up with it, for he is also thinking about what and who he wants to be (the words we hear aren't being spoken out loud by anyone, they are in the young man's head – something which the musical context makes seem natural, for when you listen to music you seem to leave the world of people and things, and enter one of thought and feeling. Or at least, that is *one* of the many experiences that music has to offer.)

Later in the commercial the young man appears as a musician. There is one episode where he is playing with his band, backed by two attractive girls (Fig. 2). Everything is lurex and sequins; this is glamour, this is the real thing, this is what being a musician is all about . . . But the sequence is no more than a fantasy (you can tell this because, unlike the rest of the commercial, it is shot in black and white), and the picture dissolves into a scene in a shopping mall – Whitley's shopping centre in Bayswater, to be precise (Fig. 3). The young man is still there, but his electronic keyboard has turned into a piano – and the pretty girls have

1

Sample of neutral reading material

A.10 List of questions for mathematical task

1 **Question 1:**
401 - 267

2 **Question 2:**
460 ÷ 8

3 **Question 3:**
43 x 36

4 **Question 4:**
32.5 x 9

5 **Question 5:**
4.8 x 3.4

6 **Question 6:**
4t - 52 = 588

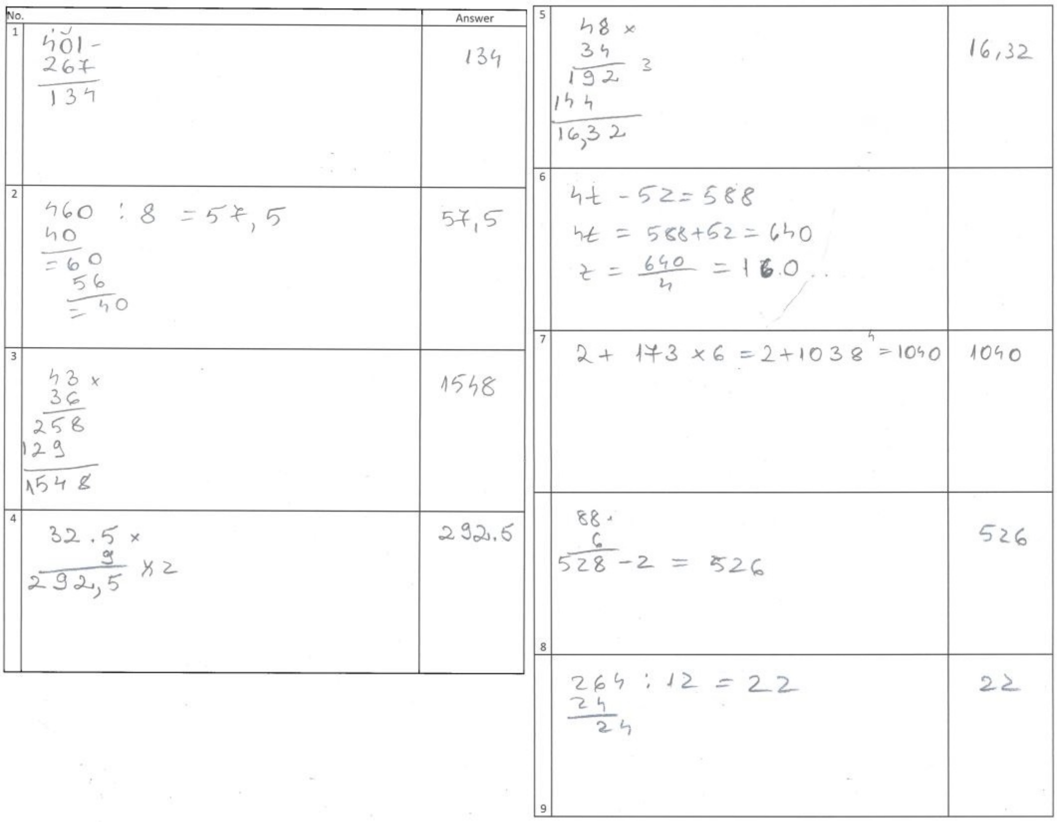
7 **Question 7:**
2 + 173 x 6

8 **Question 8:**
88 x 6 - 2

9 **Question 9:**
264 ÷ 12

List of questions for mathematical task.

A.11 Calculation sheet



No.		Answer
1	401 - 267 = 134	134
2	460 : 8 = 57,5	57,5
3	43 × 36 = 1548	1548
4	32.5 × 9 = 292,5 ×2	292.5
5	48 × 34 = 16,32	16,32
6	4t - 52 = 588; 4t = 588+52 = 640; t = 640/4 = 160	
7	2 + 173 × 6 = 2+1038 = 1040	1040
8	88 × 6 = 528 - 2 = 526	526
9	264 : 12 = 22	22

Sample of calculation sheet.

A.12 Math anxiety questionnaire

Please respond to each of the following statements.

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
Math makes me feel comfortable and easy.					
Math is most dreaded subject for me.					
I feel worried before entering the math class.					
I find math interesting.					
Math is one of my favorite subjects.					
	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
I am always afraid of math exams.					
Solving math problems is always pleasant for me.					
I feel nervous when I am about to do math homework.					
I feel happy and excited in a math class as compared to any other class.					
I would prefer math as one of my subjects in higher studies.					
	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
Math is a headache for me.					
I am afraid to ask questions in math class.					
Math doesn't scare me at all.					
My mind goes blank when teacher asks math questions					

Math anxiety questionnaire.

A.13 The Perceived Stress Scale questionnaire

The questions in this scale ask you about your feelings and thoughts **during the last month**. In each case, please indicate how often you felt or thought a certain way.

	Never	Almost never	Sometimes	Fairly often	Very often
How often have you been upset because of something that happened unexpectedly?					
How often have you felt that you were unable to control the important things in your life?					
How often have you felt nervous and "stressed"?					
How often have you felt confident about your ability to handle your personal problems?					
How often have you felt that things were going your way?					
	Never	Almost never	Sometimes	Fairly often	Very often
How often have you found that you could not cope with all the things that you had to do?					
How often have you been able to control irritations in your life?					
How often have you felt that you were on top of things?					
How often have you been angered because of things that were outside of your control?					
How often have you felt difficulties were piling up so high that you could not overcome them?					

The Perceived Stress Scale (PSS-10) questionnaire.

A.14 Questions related to emotions while interact with Paro

Please rate your emotions when you were playing with Paro.
(1 = not at all, 7 = extremely)

	1	2	3	4	5	6	7
Bored	○	○	○	○	○	○	○
Relaxed	○	○	○	○	○	○	○
Disgusted	○	○	○	○	○	○	○
Excited	○	○	○	○	○	○	○
Confused	○	○	○	○	○	○	○
Happy	○	○	○	○	○	○	○

Emotions questionnaire