DISS. ETH NO. 27731

HUGGIEBOT: AN INTERACTIVE HUGGING ROBOT WITH VISUAL AND HAPTIC PERCEPTION

A dissertation submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

ALEXIS E. BLOCK
M.S.E. in Robotics
University of Pennsylvania

born on 23 December 1993
citizen of the United States of America

accepted on the recommendation of

Prof. Dr. O. Hilliges, examiner
Prof. Dr. K. J. Kuchenbecker, co-examiner
Prof. Dr. M. Mataric, co-examiner

2021

Alexis E. Block: *HuggieBot: An Interactive Hugging Robot with Visual and Haptic Perception*, © 2021

DOI: 10.–/ethz-a-

To Daddy,
I wish more than anything that I could hug you one more time.

ABSTRACT

Hugs are one of the first forms of contact and affection humans experience. Receiving a hug is one of the best ways to feel socially supported, and the lack of social touch can have severe adverse effects on an individual's well-being. Due to the prevalence and health benefits of hugging, roboticists are interested in creating robots that can hug humans as seamlessly as humans hug other humans. However, hugs are complex affective interactions that need to adapt to the height, body shape, and preferences of the hugging partner, and they often include intra-hug gestures like squeezes. This dissertation aims to create a series of hugging robots that use visual and haptic perception to provide enjoyable interactive hugs. Each of the four presented HuggieBot versions is evaluated by measuring how users emotionally and behaviorally respond to hugging it; HuggieBot 4.0 is explicitly compared to a human hugging partner using physiological measures.

Building on research both within and outside of human-robot interaction (HRI), this thesis proposes eleven tenets of natural and enjoyable robotic hugging. These tenets were iteratively crafted through a design process combining user feedback and experimenter observation, and they were evaluated through user studies. A good hugging robot should (1) be soft, (2) be warm, (3) be human-sized, (4) autonomously invite the user for a hug when it detects someone in its personal space, and then it should wait for the user to begin walking toward it before closing its arms to ensure a consensual and synchronous hugging experience. It should also (5) adjust its embrace to the user's size and position, (6) reliably release when the user wants to end the hug, and (7) perceive the user's height and adapt its arm positions accordingly to comfortably fit around the user at appropriate body locations. Finally, a hugging robot should (8) accurately detect and classify gestures applied to its torso in real time, regardless of the user's hand placement, (9) respond quickly to their intra-hug gestures, (10) adopt a gesture paradigm that blends user preferences with slight variety and spontaneity, and (11) occasionally provide unprompted, proactive affective social touch to the user through intra-hug gestures. We believe these eleven tenets are essential to delivering high-quality robot hugs. Their presence results in a hug that pleases the user, and their absence results in a hug that is likely to be inadequate. We present these tenets as guidelines for future

hugging robot creators to follow when designing new hugging robots to ensure user acceptance.

We tested the four versions of HuggieBot through six user studies. First, we analyzed data collected in a previous study with a modified Willow Garage Personal Robot 2 (PR2) to evaluate human responses to different robot physical characteristics and hugging behaviors. Participants experienced and evaluated twelve hugs with the robot, divided into three randomly ordered trials that focused on physical robot characteristics (single factor, three levels) and nine randomly ordered trials with low, medium, and high hug pressure and duration (two factors, three levels each).

Second, we created an entirely new robotic platform, HuggieBot 2.0, according to our first six tenets. The new platform features a soft, warm, inflated body (HuggieChest) and uses visual and haptic sensing to deliver closed-loop hugging. We first verified the outward appeal of this platform compared to the previous PR2-based HuggieBot 1.0 via an online video-watching study involving 117 users. We then conducted an in-person experiment in which 32 users each exchanged eight hugs with HuggieBot 2.0, experiencing all combinations of visual hug initiation, haptic sizing, and haptic releasing.

We then refine the original fourth tenet (visually perceive its user) and present the remaining five tenets for designing interactive hugging robots; we validate the full list of eleven tenets through more in-person studies with our custom robot. To enable perceptive and pleasing autonomous robot behavior, we investigated robot responses to four human intra-hug gestures: holding, rubbing, patting, and squeezing. The robot's inflated torso's microphone and pressure sensor collected data of 32 people repeatedly demonstrating these gestures, which were used to develop a perceptual algorithm that classifies user actions with 88% accuracy. From user preferences, we created a probabilistic behavior algorithm that chooses robot responses in real time. We implemented improvements to the robot platform to create a third version of our robot, HuggieBot 3.0. We then validated its gesture perception system and behavior algorithm in a fifth user study with 16 users.

Finally, we refined the quality and comfort of the embrace by adjusting the joint torques and joint angles of the closed pose position, we further improved the robot's visual perception to detect changes in user approach, we upgraded the robot's response to users who do not press on its back, and we had the robot respond to all intra-hug gestures with squeezes to create our final version of the robotic platform, HuggieBot 4.0. In our sixth user study, we investigated the emotional and physiological effects of hugging a robot compared to the effects of hugging a friendly but unfamiliar person. We continuously monitored participant heart rate and collected saliva samples at seven time points across the 3.5-hour study to measure the temporal evolution of cortisol and oxytocin. We used an adapted Trier Social Stress Test (TSST) protocol to reliably and ethically induce stress in the participants. They then experienced one of five different hug intervention methods before all interacting with HuggieBot 4.0.

The results of these six user studies validated our eleven hugging tenets and informed the iterative design of HuggieBot. We see that users enjoy robot softness, robot warmth, and being physically squeezed by the robot. Users dislike being released too soon from a hug and equally dislike being held by the robot for too long. Adding haptic reactivity definitively improves user perception of a hugging robot; the robot's responses and proactive intra-hug gestures were greatly enjoyed. In our last study, we learned that HuggieBot can positively affect users on a physiological level and is somewhat comparable to hugging a person. Participants have more favorable opinions about hugging robots after prolonged interaction with HuggieBot in all of our research studies.

RÉSUMÉ

Les câlins sont l'une des premières formes de contact et d'affection que connaissent les humains. Recevoir un câlin est l'un des meilleurs moyens de se sentir soutenu socialement, et l'absence de contact social peut avoir de graves effets négatifs sur le bien-être d'une personne. En raison de la fréquence et des avantages des câlins pour la santé, les roboticiens s'intéressent à la création de robots capables de faire des câlins à des humains, aussi facilement que les humains font des câlins à d'autres humains. Cependant, les câlins sont des interactions affectives complexes qui doivent s'adapter à la taille, à la morphologie et aux préférences de la personne qui les fait, et ils incluent souvent des gestes « intra-câlin » comme les serrages. Cette thèse vise à créer une série de robots câlineurs qui utilisent la perception visuelle et haptique pour fournir des câlins interactifs agréables. Chacune des quatre versions de HuggieBot présentées est évaluée en mesurant la façon dont les utilisateurs répondent émotionnellement et au niveau comportemental à ses câlins ; HuggieBot 4.0 est explicitement comparé à un partenaire humain à l'aide de mesures physiologiques.

S'appuyant sur des recherches menées à l'intérieur et à l'extérieur du domaine d'étude « interactions homme-robot » (IHB), cette thèse propose onze principes (les "commandements") d'un câlin robotique naturel et agréable. Un bon robot câlin doit (1) être doux, (2) être chaud, (3) être de taille humaine, (4) inviter de manière autonome l'utilisateur à un câlin lorsqu'il détecte quelqu'un dans son espace personnel, puis il doit attendre que l'utilisateur commence à marcher vers lui avant de fermer ses bras pour assurer une expérience de câlin consensuelle et synchrone. Il doit également (5) adapter son câlin à la taille et à la position de l'utilisateur, (6) relâcher de manière fiable lorsque l'utilisateur souhaite mettre fin au câlin, et (7) percevoir la taille de l'utilisateur et adapter la position de ses bras en conséquence pour s'adapter confortablement à l'utilisateur aux endroits appropriés du corps. Enfin, un robot câlin doit (8) détecter et classifier avec précision les gestes appliqués à son torse en temps réel, quel que soit le placement de la main de l'utilisateur, (9) répondre rapidement auxgestes de câlin, (10) adopter un paradigme gestuel qui combine les préférences de l'utilisateur avec une légère variété et spontanéité, et (11) fournir occasionnellement un contact social affectif proactif et spontané à l'utilisateur par des gestes de câlin. Nous pensons que ces onze principes

sont essentiels pour fournir des câlins robotisés de haute qualité. Leur présence donne lieu à un câlin qui plaît à l'utilisateur, et leur absence donne lieu à un câlin qui risque d'être inadéquat. Nous présentons ces commandements comme des lignes directrices que les futurs créateurs de robots câlineurs devront suivre lors de la conception de nouveaux robots câlineurs afin de garantir l'adoption par les utilisateurs.

Nous avons testé les quatre versions de HuggieBot à travers six expériences d'utilisateurs. Tout d'abord, nous avons analysé les données recueillies dans une étude précédente avec un robot personnel Willow Garage 2 (PR2) modifié pour évaluer les réponses humaines aux différentes caractéristiques physiques du robot et aux comportements de câlin. Les participants ont fait l'expérience et ont évalué douze câlins avec le robot, divisés en trois essais aléatoirement classés et axés sur les caractéristiques physiques du robot (facteur unique, trois niveaux), et neuf essais aléatoirement classés avec une pression et une durée de câlin faibles, moyennes et élevées (deux facteurs, trois niveaux chacun).

Deuxièmement, nous avons créé une plateforme robotique entièrement nouvelle, HuggieBot 2.0, conformément à nos six premiers principes. La nouvelle plateforme est dotée d'un corps doux, chaud et gonflé (HuggieChest) et utilise la détection visuelle et haptique pour fournir des câlins en boucle fermée. Nous avons d'abord vérifié l'attrait extérieur de cette plateforme par rapport à l'ancien HuggieBot 1.0 basé sur le PR2, par le biais d'une étude par vidéo en ligne impliquant 117 utilisateurs. Nous avons ensuite mené une expérience en personne dans laquelle 32 utilisateurs ont chacun échangé huit câlins avec HuggieBot 2.0, en expérimentant toutes les combinaisons de déclenchement visuel du câlin, de dimensionnement haptique et de relâchement haptique.

Nous affinons ensuite le quatrième principe original (percevoir visuellement l'utilisateur) et présentons les cinq autres principes pour la conception de robots câlineurs interactifs ; nous validons la liste complète des onze principes par d'autres études en personne avec notre robot personnalisé. Pour permettre un comportement autonome perceptif et agréable du robot, nous avons étudié les réponses du robot à quatre gestes humains intra-câlin : tenir, frotter, tapoter et serrer. Le microphone et le capteur de pression du torse gonflé du robot ont recueilli les données de 32 personnes montrant ces gestes de manière répétée, et qui ont été utilisées pour développer un algorithme perceptif qui classifie les actions des utilisateurs avec une précision de 88%. À partir des préférences des utilisateurs, nous avons créé un algorithme de comportement probabiliste qui choisit les réponses du robot en temps réel. Nous avons apporté des améliorations à la plateforme du robot pour créer une troisième version de notre robot, HuggieBot 3.0. Nous avons ensuite validé son système de perception des gestes et son algorithme de comportement dans une cinquième étude auprès de 16 utilisateurs.

Enfin, nous avons affiné la qualité et le confort du câlin en ajustant les couples et les angles des articulations de la position fermée, nous avons davantage amélioré la perception visuelle du robot pour détecter les changements d'approche de l'utilisateur, nous avons amélioré la réponse du robot aux utilisateurs qui n'appuient pas sur son dos, et nous avons fait en sorte que le robot réponde à tous les gestes d'étreinte par des pressions pour créer notre version finale de la plate-forme robotique, HuggieBot 4.0. Dans notre sixième étude sur les utilisateurs, nous avons étudié les effets émotionnels et physiologiques d'un câlin avec un robot par rapport aux effets d'un câlin avec une personne inconnue et amicale. Nous avons surveillé en continu la fréquence cardiaque des participants et recueilli des échantillons de salive à sept moments au cours de l'étude de 3,5 heures, afin de mesurer l'évolution temporelle du cortisol et de l'ocytocine. Nous avons utilisé un protocole adapté du Trier Social Stress Test (TSST) pour provoquer du stress de manière fiable et éthique chez les participants. Ils ont ensuite expérimenté l'une des cinq méthodes différentes de déclenchement d'un câlin avant de tous interagir avec HuggieBot 4.0.

Les résultats de ces six études d'utilisateurs ont permis de valider nos onze principes de câlins et de guider la conception itérative de HuggieBot. Nous constatons que les utilisateurs apprécient la douceur du robot, la chaleur du robot et le fait d'être serré physiquement par le robot. Les utilisateurs n'aiment pas être libérés trop tôt d'un câlin et n'aiment pas non plus être tenus par le robot pendant trop longtemps. L'ajout de la réactivité haptique améliore définitivement la perception de l'utilisateur d'un robot câlineur ; les réponses du robot et les gestes intra-câlin ont été très appréciés. Dans notre dernière étude, nous avons appris que HuggieBot peut avoir un effet positif sur les utilisateurs à un niveau physiologique et qu'il est quelque peu comparable à un câlin avec une personne. Les participants ont des opinions plus favorables sur les robots câlineurs après une interaction prolongée avec HuggieBot dans toutes nos études de recherche.

ACKNOWLEDGMENTS

This work is supported in part by funding from the Max Planck ETH Center for Learning Systems, the Max Planck Institute for Intelligent Systems, ETH Zürich, and the IEEE Technical Committee on Haptics (TCH).

First of all, I would like to thank my advisor, Prof. Dr. Katherine J. Kuchenbecker, for her advice and guidance over the last nine years. I am grateful for the opportunity to join her in Germany for my doctorate. My Ph.D. was a great learning experience, and I am thankful for her patience and trust in me. Furthermore, I greatly appreciate the countless opportunities she provided by supporting me and my research - like attending conferences, summer schools, and speaking publicly about our joint work.

I am also extremely grateful for my two ETH advisors, Prof. Dr. Otmar Hilliges and Prof. Dr. Roger Gassert. I greatly appreciate their willingness to join this collaborative project and share our vision for providing people with high-quality robot hugs. I am especially thankful to Otmar for agreeing to work with me, despite my not having the necessary relevant background for his group. I learned so much from working with him and the members of his lab. I appreciate how kind and welcoming both Otmar and Roger's labs were to having me join them for a year and a half in the middle of my doctoral studies.

I want to thank Kinova Robotics for supporting various aspects of this research project, and especially for providing us with an emergency replacement JACO arm right before a big user study.

I also thank Elyse Chase for developing the artistic rendering of our project goal (Fig. 1.1), as well as Elisabeth Smela and Ying Chen for providing us with custom-made sensors for HuggieBot 1.0. I thank Joe Romano, whose PR2 props code served as a foundation of this experiment, Siyao (Nick) Hu for his help and expertise working with the PR2, and Naomi Fitter for her help in designing the user study and analysis for HuggieBot 1.0.

I am grateful for our co-authors, Hasti Seifi, Sammy Christen, and Shari Young Kuchenbecker, not only for their contributions that merit co-authorship, but also for their constant support and guidance throughout the project.

I would like to thank Bernard Javot for his help in making the base, stand, and head of HuggieBot 2.0.

I thank Michaela Wieland and Meike Pech for their help in organizing the transport of HuggieBot from Germany to Switzerland and back. I also thank Ilona Jacobi for assisting with this transportation process, all logistical arrangements involving HuggieBot, and her kindness, patience, and help with German translations.

I appreciate the IT support from Joey Burns and Felix Grüninger when problems arose just before or during an experiment.

I thank Ravali Gourishetti for her careful review of our user studies that fell within the Haptic Intelligence Framework Agreement with the Max Planck Ethics Council.

I greatly appreciate the help from Nati Egana, who helped on countless occasions to turn on and test the robot when Alexis was working remotely. She also was instrumental in our final study by preparing all participant folders and doing laundry to ensure clean, hygienic robot and human hugging outfits each day. She also went through participant videos and uploaded the necessary segments to make video coding easier. I sincerely thank her for her hard work and dedication to our project.

I am thankful for the help from Lijuan Wang for arranging and rearranging the delivery and pickup of the freezer to store the saliva samples. I also appreciate her dedication and diligence with the data entry for all the surveys from our last study.

I sincerely thank the members of the HuggieBot Theatre Troupe: Silvia Passera, Manuel Krstanovic, Uwe-Peter Spinner, Claudia Gallatz, and Dorothea Baltzer, for their hard work and dedication. Running the last user study in Chapter 7 would not have been possible without their help.

I am also appreciative to the entire Haptic Intelligence Department, who have all contributed at some point to various aspects of supporting this project. In particular, I would like to point out the help from Ben Richardson, Mayumi Mohan, Rachael Bevill Burns, Natalia Sanchez-Tamayo, Neha Thomas, Maria-Paola Forte, and Valerio Ortenzi for providing feedback on the quality of the embrace, listening to de-bugging explanations, locating holes and helping reseal them in popped HuggieChests, pilot testing experiments, and general support and enthusiasm during long experiment days and weeks.

A very special thanks goes to my family and friends who supported the emotional roller coaster that is a Ph.D. Thank you Mommy, Joshie, and Nanny. Thank you to my husband, Samuel Bitter, and his parents, Lydia and Simon Bitter. Thank you Elizabeth Silvestro, Sean Lynch, Shaanan Cohney, Melanie Schuler, Michael Dea, and Liz Walton. I love you all very much.

CONTENTS

1

INTRODUCTION

We need 4 hugs a day for survival. We need 8 hugs a day for maintenance. We need 12 hugs a day for growth.

— Virginia Satir, Family Therapy Pioneer

1.1 MOTIVATION

From the moment we are born, social touch affects our future ability to function well in society. Infants who are held by their mothers for two hours after they are born have better interactions with their mothers and are better at handling stress (Uvnäs-Moberg, Handlin, & Petersson, 2014). In such a close, positive relationship, the hormone oxytocin is released when the two partners see, hear, or even think of each other. In turn, this release bonds them even more closely and improves positive human relationships. As a hug is one of the first forms of and most common kinds of contact many people receive, we were interested in investigating whether its benefits could be derived when the hugging partner is a robot rather than another human.

1.1.1 *Physiological Effects of Human Hugs*

Researchers have looked into the social and physical health benefits of hugs, often by studying the body's positive physiological response to human hugs. Deep pressure touch therapy is the kind of touch we receive from someone touching or hugging us firmly (Edelson et al., 1999). While the easiest and perhaps most common way to receive this beneficial touch is from another person, over time as we grow, rather than relying on a parent for comfort, humans learn to self-soothe in similar ways, such as holding themselves, rubbing their arms, and wrapping themselves tightly in a warm blanket (Uvnäs-Moberg, Handlin, & Petersson, 2014). These methods work because they are reminiscent of the earliest form of deep pressure touch they received. Even when this touch comes from the user's own actions, it still has a calming effect and has been shown to alleviate stress and anxiety dramatically and lower heart rates and blood pressure (Edelson et al., 1999). This beneficial touch results in an oxytocin increase, which produces strong feelings of social bonding (Uvnäs-Moberg, Handlin, & Petersson, 2014).

Many researchers seeking to understand this cascade have studied the first form of affective, social touch we receive: a mother holding and soothing her newborn child. Uvnäs-Moberg, Handlin, and Petersson (2014) found the embrace between mother and child to cause a powerful oxytocin response. In addition to calming the child down, this helpful touch (which can come in the form of cradling, hugging, kissing, stroking, etc.) strengthens the bond between parent and child and increases the body's production of oxytocin in both. An increase in oxytocin provides a host of benefits, including a greater tolerance for pain and stress (Ditzen, Neumann, Bodenmann, von Dawans, et al., 2007b). Also, Uvnäs-Moberg, Handlin, and Petersson (2014) note that it can have a lasting effect on our ability to connect emotionally with people throughout our lives. The absence of social touch can have detrimental effects on child development (Cascio, Moore, & McGlone, 2019). Infants who do not receive this beneficial touch at birth will have trouble forming relationships as they grow.

One group of researchers found that warm partner contact increases oxytocin levels for both men and women. However, the importance of oxytocin and its effects on blood pressure may be greater for women (Grewen et al., 2005). Another group of researchers found that women who received positive physical partner contact before stress exhibited lower cortisol levels and heart rate than those who did not, while verbal social support alone did not help in reducing stress (Ditzen, Neumann, Bodenmann, von Dawans, et al., 2007a). Light, Grewen, and Amico (2005) found that frequent hugs between long-term partners resulted in lower blood pressure and higher oxytocin levels in premenopausal women.

People interact with each other through various kinds of social touch that change across the lifetime in order to foster a sense of community, strengthen relationships, and build trust (Cascio, Moore, & McGlone, 2019; Cohen et al., 2015). Social touch in a broader sense is also vital for maintaining many kinds of relationships among humans and primates alike (Suvilehto et al., 2015); hugs seem to be a basic evolutionary need. They are therefore highly popular! An online study conducted in 2020 polled 1,204,986 people to find out "what is the best thing?" Hugs earned fifth place out of 8,850 things, behind only sleep, electricity, the Earth's magnetic field, and gravity (Scott, 2020).

1.1.2 *Physical Properties of Social Touch*

Physical properties of objects strongly affect how contact interactions are perceived. Harlow and Zimmermann's work with infant Rhesus monkeys (Harlow & Zimmerman, 1959) strongly influenced this thesis. When given the choice of which surrogate mother they preferred, a wire mother who fed them, or a cloth mother who did not, overwhelmingly the infants chose the cloth mother. A similar phenomenon can be observed in human children, who prefer to sleep with plush comfort objects or blankets because of the emotional attachment they develop to the experience of softness (Mahalski, 1981). These studies inspired our interest in robot softness.

Another comforting physical property that often accompanies softness is warmth. Williams and Bargh ran a set of experiments on this topic; in Study 1, users briefly held a cup of warm or iced coffee and were then asked to judge a target person's personality (Williams & Bargh, 2008). Participants who held the warm cup of coffee associated warmer personality traits (generous, caring, etc.) to the target person than their cold coffee counterparts. Note that replication of the results of Study 2, which measured a person's subsequent tendency toward prosocial behavior, was not upheld by Lynott et al. (Lynott et al., 2014). To the best of our knowledge, researchers have not had difficulty replicating the Study 1 finding that humans associate social warmth with warm physical contact. Indeed, Lakoff and Johnson have suggested that these two experiences go hand in hand (Lakoff & Johnson, 1999): because we receive them simultaneously so often as children, when we experience physical warmth (a warm hug or being wrapped in a blanket), we associate it with feelings of social warmth (love). These non-robotic studies inspired the integration of heat into our robotic platform.

1.1.3 *Social Touch Between Humans*

As humans age further, researchers have found that the areas a person is allowed to touch on another's body are directly correlated with the strength of the relationship between the two people (Suvilehto et al., 2015). The closer two people are emotionally to each other, the more areas they are allowed to contact for social touch. Appropriate location of touch was of particular importance to our research, as we found users to be highly sensitive to the hand placement of our hugging robot; placing the robot's arms either too low or too high on the body was not appreciated.

In addition to the location of social touch, another critical element to consider during social touch between humans is the applied contact strength. Too little pressure can sometimes create a disingenuous impression or not provide enough emotional support, while too much pressure can hurt your partner. Researchers studied the physiological responses (heart rate and R-R interval) of infants being held with different levels of tightness and by people with varying relationships to the child (Yoshida et al., 2020). Infants responded best when being held with a medium amount of contact pressure by a parent. This study identifies a Goldilocks zone where the pressure is not too low and not too high, but "just right." This pressure zone may vary across people and can even change for a single person depending on how they feel and how much support they currently need (more support may require more applied pressure).

As previously mentioned, a common way to appropriately provide a person with deep pressure touch is through a hug, where both participants wrap their arms around the other person's torso. During prolonged hugs (lasting more than three seconds), the two hugging partners rarely remain stationary in the embrace (Nagy, 2011). In both humans and primates alike, common intra-hug gestures include stroking/rubbing, patting, and squeezing (de Waal, 2019). These gestures provide comfort and help the receiver feel more emotionally supported than if the partner simply stood still during the embrace. Although few other investigations have explored intra-hug gestures, this research inspired us to create a robot that can detect and respond to such gestures during a hug to more closely mimic humans and primates so that it can eventually, hopefully, provide better emotional support to its users.

1.1.4 *Summary*

With the health benefits and prevalence of hugs in daily human interactions, it is natural that roboticists have tried to create this gesture artificially. A major challenge of mechanizing hugs is the safety and comfort of the human during this intimate exchange. Researchers, therefore, have taken many different approaches, as summarized in Chapters 2 and 3.

This thesis investigates the design and evaluation of several versions of the first interactive human-sized hugging robot with visual and haptic perception. In particular, we present a novel, custom robot that autonomously adjusts to user height, size, approach, desired hug duration, and intra-hug gestures.

We conduct and discuss six user studies that evaluate four iterative versions of our robotic platform, HuggieBot, over Chapters 4 through 7. Overall, our work has demonstrated that an autonomous, interactive hugging robot is a promising new technology from which many users could gain enjoyment and health benefits.

1.2 PROBLEM SETTING

Not everyone is fortunate enough to have close, positive relationships with people around them. If a robotic embrace can achieve a calming effect similar to that caused by a hug from another person, perhaps this helpful supplementary touch can benefit people who otherwise would not be able to experience hugs. The broader goal of this thesis is to *create an embodied affective robot that can provide beneficial hugs in situations when requesting this form of comfort from other humans is difficult or impossible*.

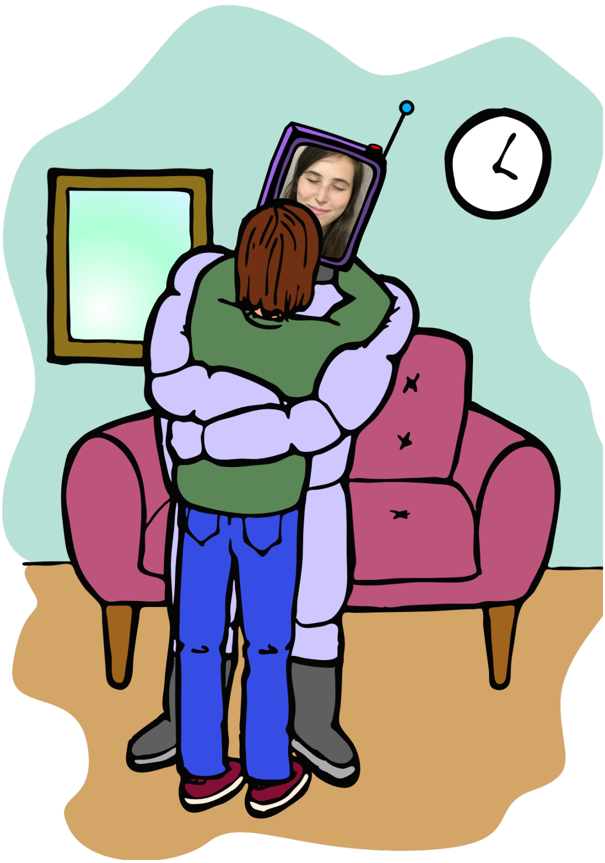
Many common examples where people lack access to human hugs stem from long-term physical separation. Examples of long-term physical separation may be a parent who frequently travels for work, a student studying in another city, state, or country for university, a senior in a nursing home, or a patient with extended stays in hospitals.

Unfortunately, ever more interactions are happening remotely and online, especially during this unprecedented time of physical distancing due to COVID-19. During the current global pandemic, some family members have been unable to come into close contact with each other for more than one year, and the effects are showing. An increasing number of people are suffering from a lack of social touch, which can be detrimental to both our physical and mental health (Cascio, Moore, & McGlone, 2019; Morrison & Gore, 2010; Neira & Barber, 2014).

Our long-term research goal is to determine the extent to which we can strengthen personal relationships that are separated by a physical distance via hugging robots that provide high-quality social touch. Figure 1.1 shows an artist's rendering of this long-term goal. We seek to supplement, not replace, human touch.

1.3 CONTRIBUTIONS

In this thesis, we make the following contributions towards enabling pleasant, close, personal, social-physical human-robot interactions:



.

FIGURE 1.1: An artistic rendering of the end goal of this thesis: a social humanoid robot that can capably hug humans in everyday settings. The screen shows a video message from the person who sent a customized hug to the human participant. Further discussion of the situation depicted in this image can be found in Chapter 8

ADAPTIVE HUGGING AS A FORM OF GRASPING Our first contribution is our unique approach to adaptive hugging. People of every shape and size deserve to receive a high-quality embrace that adapts to their body position. As grasping objects of different sizes and shapes is a common and well-studied problem in robotics, we, therefore, model hugging after a two-fingered gripper. This method ensures a secure, custom, and comfortable embrace for each individual user that neither leaves air gaps nor applies excessive pressure to the user's body.

HUGGIECHEST A second contribution is the creation of the HuggieChest. HuggieChest is a novel, inflatable torso for a human-sized hugging robot that simultaneously softens and provides sensing capabilities. This torso can be used in real time to detect both coarse and fine contacts made on the robot's torso. We define coarse contacts as significant changes in pressure, such as making or breaking contact with the torso, which we use to determine when a user had started or stopped hugging the robot. We also developed a detection and classification algorithm using the two data streams from the HuggieChest (a pressure sensor and a microphone) to determine in real-time which fine contacts, or intra-hug gestures (rubs, pats, and squeezes), the user is performing on the robot's torso. The creation of this torso is a major contribution to our development of the first interactive hugging robot.

HUGGIEBOT A crucial contribution of this thesis, of course, is the HuggieBot robotic platform. It is an autonomous, interactive, fully integrated robot with visual and haptic perception. All components of this robot were custom-made, except for two Kinova JACO arms. All software for controlling the robot's arms, torso, face, voice, and camera was also custom-written for the robot. We created a custom v-shape based to allow users to easily get close to the robot without having to lean over a large base (as they did with the PR2). The stand is height adjustable if desired, though for all our studies, we kept the robot height at 1.57 m. The HuggieChest makes up the torso of the robot. On top of the torso are heating pads, a robe, and a sweatshirt. The Kinova JACO end effects are covered in mittens. We designed and 3D printed a head for the robot, which houses the robot's face screen, computer, speaker, camera, and many cables. We created several animated faces that cycle through different smiles to appear more "alive." All a user has to do is walk in front of the robot, and it will autonomously detect them and invite them for a hug. It will remain with its arms outstretched until the user walks forward, at which point the robot will provide them with a custom adaptive embrace that lasts as long as the user wants.

ELEVEN HUG TENETS Our final contribution is the eleven tenets of pleasant and enjoyable robotic hugging. We recommend that future hugging robot designers follow these tenets to optimize the user experience.

A hugging robot should:

1. be soft,
2. be warm,
3. be human-sized,
4. autonomously invite the user for a hug when it detects someone in its personal space, and then it should wait for the user to begin walking toward it before closing its arms to ensure a consensual and synchronous hugging experience,
5. adjust its embrace to the user's size and position,
6. reliably release when the user wants to end the hug,
7. perceive the user's height and adapt its arm positions accordingly to comfortably fit around the user at appropriate body locations,
8. accurately detect and classify gestures applied to its torso in real time, regardless of the user's hand placement
9. respond quickly to their intra-hug gestures,
10. adopt a gesture response paradigm that blends user preferences with slight variety and spontaneity,
11. and occasionally provide unprompted, proactive affective social touch to the user through intra-hug gestures.

SPECIAL NOTE ON HUGGIEBOT 1.0 The custom hardware and software upgrades to the Willow Garage Personal Robot 2 (PR2) and the user study were conducted as part of my master's thesis while at the University of Pennsylvania (Block, 2017). All analyses and writing were conducted during my doctoral studies. The associated publication was submitted, revised, and ultimately accepted during my first two years of doctoral research (Block & Kuchenbecker, 2019). Because my doctoral thesis heavily builds on the data

I collected in my master's thesis research, I have included HuggieBot 1.0 in my doctoral dissertation, as found in Chapter 4.

1.4 STRUCTURE OF THESIS

The structure of this dissertation will be as follows: after establishing the background knowledge and state-of-the-art on social-physical human-robot interaction, we outline our main contributions. More specifically:

CHAPTER 2 discusses the background of technology-mediated social touch for uses other than hugging.

CHAPTER 3 provides an overview of the state of the art of robots for social-physical human-robot interaction, with specific focus on other hugging robots.

CHAPTER 4 presents the design and evaluation of the first version of HuggieBot. This chapter is based on our 2019 publication in the International Journal of Social Robotics (Block & Kuchenbecker, 2019).

CHAPTER 5 proposes the second version of HuggieBot, along with the "Six Hug Tenets" of pleasant and enjoyable robotic hugging. This chapter is based on our 2018 publication in the Companion Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (Block & Kuchenbecker, 2018) and our 2021 paper in the Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (Block, Christen, et al., 2021).

CHAPTER 6 further refines HuggieBot into a third version with significant software improvements enabling a more interactive experience, and expands the prior "Six Hug Tenets" to "Eleven Hug Tenets." Two user studies are discussed which helped create and test the detection and classification algorithm and probabilistic behavioral algorithm to enable the robot to respond to user intra-hug gestures. This chapter is based on our manuscript which is presently under review for possible publication in the Special Issue on Designing the Robot Body: Critical Perspectives on Affective Embodied Interaction of the ACM Transactions on Human-Robot Interaction (Block, Seifi, et al., 2021).

CHAPTER 7 presents a fourth version of HuggieBot which was used in a large-scale user study to measure the physiological effects of robot hugs and compare them to those of hugging a person. This chapter is based on a manuscript that is currently in preparation for submission (Block, Young Kuchenbecker, et al., 2021).

CHAPTER 8 summarizes and discusses our findings and provides an overview of future directions of research on social-physical human-robot interaction to support high-quality interactive robot hugs for all users.

1.5 PUBLICATIONS

The contributions of this thesis are based on the following publications:

Block, Alexis E, Christen, S., Gassert, R., Hilliges, O., & Kuchenbecker, K. J. (2021). The six hug commandments: Design and evaluation of a human-sized hugging robot with visual and haptic perception. *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 380.

Block, Alexis E, & Kuchenbecker, K. J. (2018). Emotionally supporting humans through robot hugs. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 293.

Block, Alexis E, & Kuchenbecker, K. J. (2019). Softness, warmth, and responsiveness improve robot hugs. *International Journal of Social Robotics*, *11*(1), 49.

Block, Alexis E, Seifi, H., Gassert, R., Hilliges, O., & Kuchenbecker, K. J. (2021). In the arms of a robot: Designing autonomous hugging robots with intra-hug gestures [under review]. *ACM Transactions on Human-Robot Interaction (THRI) Special Issue on Designing the Robot Body: Critical Perspectives on Affective Embodied Interaction*, 1.

Block, Alexis E, Young Kuchenbecker, S., Gassert, R., Hilliges, O., & Kuchenbecker, K. J. (2021). Love, actually? evaluating the emotional and physiological effects of robot hugs [in preparation]. *Journal to be determined*.

2 BACKGROUND

I have learned that there is more power in a good strong hug than in a thousand meaningful words.
— Ann Hood, American Novelist

Social interactions provide the foundation for culture and society. These interactions can occur in different forms, some via physical contact and some without touch. Historically, all kinds of social interactions took place in person. Then gradually, other mediums enabled these interactions to become physically distanced - like letter writing, phone calls, and texting. Now, technology is providing more opportunities for social interactions, both with and without physical touch.

This chapter provides a brief review to familiarize the reader with the various technology solutions to different social interactions, which have greatly inspired this thesis. First, we discuss the robotic solutions to non-touch social interactions. Next, we delve into non-robotic technology-mediated social touch solutions, which we split into two groups: wearables and comfort objects. Finally, we discuss robots designed for social touch, divided into two groups: the tactile experience of robotic touch, and robot touch for social enjoyment.

2.1 ROBOTIC NON-TOUCH SOCIAL INTERACTION

Improving the sociability of robots has interested many researchers. Most often, robots do not physically interact with people, but rather they use verbal or non-verbal cues to achieve more natural human-robot interactions. In one study about HRI in the wild, Garrell et al. found that more non-contact human-robot interactions occurred when the robot initiated the conversation (Garrell et al., 2017). People also felt the interaction was more natural when the robot gestured or initiated engagement. Such movements led the human to perceive the robot as having a higher level of intelligence or sociability. From this research, we learn that it is important to have our robot ask for a hug and gesture to the human, to increase the naturalness of the exchange.

One such group of researchers has studied the aspects that lead to increased acceptance of robots by humans. May et al. found that when people attributed human/animal-like characteristics to the robots with whom they were interacting, the participants were more excited and comfortable during the exchange (May et al., 2017). Hence, we attempt to humanize our robot by clothing it in an outfit, rather than having it look mechanical. However, Fitter et al. (2021), found that while users enjoy and appreciate personalized tele-presence robots, they are particularly bothered by robots wearing wigs. Thus, we left our robot's head uncovered.

Glas et al. were also interested in reproducing human greetings in robots for interactions in the wild (Glas et al., 2017). By programming a robot to comment on patterns or new behaviors, these researchers gave mall visitors a delight when the robotic greeter welcomed them with increasing familiarity. As greetings take different forms in different cultures (a hello, a wave, a kiss (or two), or a hug), we build on Glas et al.'s research to humanize robot greetings for people of different cultures. We also have our hugging robot take social cues. for the hug duration and any intra-hug gestures from the user, to respect different personal and cultural preferences for hugs.

In human-human exchanges like conversations, one often attempts to match the level of enthusiasm of their partner (Bernieri & Rosenthal, 1991). This phenomenon is often referred to as "The Chameleon Effect" (Chartrand & Bargh, 1999). Anzalone et al. replicated this finding with a robot (Anzalone et al., 2017): they programmed a robot to estimate the level of extraversion of their human partner and adapt its personality/behavior to better match that of the human. The humans perceived the robot to have higher social intelligence when it matched their level of extraversion compared to when the robot did not. We apply this technique to our own research by having conditions that terminate the hug based on user input, and by having the robot respond to user intra-hug gestures with a gesture of a similar level of extraversion.

Many researchers have been primarily interested in the interaction between children and robots. Some groups have focused on what happens between children and robots in unsupervised social interactions - would the children and robots get along, play and be friendly? A group of researchers deployed a robot near a group of children in a shopping mall in Japan and found the children bullied and repeatedly verbally and physically abused the robot (Nomura et al., 2016), and their finding was not unique (Brščić et al., 2015; Salvini et al., 2010). We will focus on researchers who have looked at what happens in supervised social interactions between children and robots. Much literature exists on this topic, here we present a small sample to support the claim that children quickly adapt to and work well with robots, but it is by no means an exhaustive list. Pulido et al., for instance, found that children in particular enjoy interacting with robots and are able to use them readily without assistance (Pulido et al., 2017). They also found that children worked harder at their exercises with the aid of a robot. These results indicate that robots could provide an exciting motivational source for children. Another group of researchers found that even though children ages 3-4 years old typically have no perspective taking ability, through playing "hide and seek" with a robot, the children were able to learn features and relations of objects (Trafton et al., 2006). Another group of researchers created a fluffly baby dragon robot for storytelling to help preschool aged children with language development (Kory & Breazeal, 2014). The researchers created a backchannel opportunity predication model based on features of the child's storytelling to produce contingent backchannel responses that convey a perception of active listening that children prefer over telling story to the non-contingent robot (H. W. Park et al., 2017).

While Pulido et al. found children readily willing and able to accept robots for their own use, Hudson et al. found that when it comes to robots caring for the elderly, there is a wide range of opinions, and at times, both the youth (as seen with Nomura et al. (2016)) and elderly can be hostile towards the idea (Hudson, Orviska, & Hunady, 2017). They found that this sometimes hostile attitude was primarily due to a lack of knowledge about robots, which can be remedied by improving the robot design with enhanced safety features. Baisch et al. share a similar view in their research on user-tech fit (Baisch et al., 2017). They found that tele-presence robots can be useful for fostering social support when the user has the correct resources to properly operate the technology. Both these groups argue that a technology has a higher chance of being accepted by users when they are given instructions on how to interact safely and effectively. Again, the topic of older adults and technology has a large literature scope, and only a small subset will be selected here to support the claim that with proper training and safety instructions, hostility and negative attitudes can be avoided and the benefits for older adults can be reached. One group of researchers found 21 independently living older adults (65-93 years old) open to robots performing domestic tasks, but that they preferred the robot perform tasks related to chores, manipulating objects, and information management, while they still preferred a human to help with personal care and leisure activities. (Smarr et al., 2014). (Pino et al., 2015)

Socially assistive robotics (SAR) (Feil-Seifer & Matarić, 2005) is a field of robotics that aims to help people through social rather than physical interactions. Researchers have found that such robots, particularly when expressing empathy, can help reduce pain and distress in child patients during stressful experiences, like medical visits (Trost et al., 2020). This study found that the patients interacting with an empathetic robot had more positive opinions about the robot than patients who did not have an empathetic robot. Users in the empathetic robot group also expressed believing that the robot "had feelings." From this work, we learn that users appreciate and derive more benefits from a robot they perceive as understanding and sharing their feelings. By training our robot to detect, classify, and respond to user intra-hug gestures (rubs, pats, and squeezes), we try to make our robot appear empathetic and emotionally supportive.

2.2 TECHNOLOGY-MEDIATED SOCIAL TOUCH

Because physical contact is not always possible, some researchers look for ways to strengthen relationships between emotionally close people who are separated by a distance. To bridge this gap, they often develop technology that aims to transmit social touch. The devices they create typically work in pairs, where users can send signals to each other to let their partner know they are thinking of them. The signal the other user receives indicates their partner is sending them some contact, but it may not accurately replicate the sender's desired intent (e.g., vibration output may be used to represent a squeezing input). Such objects typically fall into two categories that we discuss in more detail below: wearables and comfort objects.

Perhaps the most well-known version of technology-mediated social touch that does not fit into either of these categories is Temple Grandin's "Squeeze Machine." This technology applies lateral deep touch pressure by squeezing a user between two foam panels (Grandin, 1992). Patients on the autism spectrum, non-autistic college students, and animals all experienced similar calming effects. The Squeeze Machine is operated by the user, who can control the applied pressure and duration of the encounter, gradually building up over time as he or she becomes more comfortable. While these artificial hug recreations lack the primary component of a second partner, they do address the importance of physical touch.

2.2.0.1 *Wearables*

One related non-robotic approach is the creation of inflatable or weighted vests and jackets to help calm children with sensory processing disorder, children with attention deficit hyperactivity disorder, and individuals with autism spectrum disorder (VandenBerg, 2001). Deep pressure touch therapy, the kind received from hugging or firmly touching, has been shown to relieve anxiety for people with these disorders (Krauss, 1987). Because they require a loud pump and air flow, inflatable garments are often obtrusive and conspicuous. Inflatable or pressurized vests can also be activated remotely by a parent or instructor at any time (Duvall et al., 2016). In this instance, the child may not understand the cause of the hug. Additionally, weighted vests must constantly be removed to alleviate the pressure and then replaced.

A handheld or wearable device that connects physically separated individuals can be an ideal solution for those in a long-distance relationship. Many researchers have invested time to develop wearable technology to help physically separated loved ones feel emotionally close to each other. A common version of this wearable technology occurs in the form of a bracelet. The Squeezy bracelet (Pakanen et al., 2014) can be paired with a mobile phone and allows the user to receive haptic messages. The Hey Bracelet ("Hey Bracelet", 2020) works in pairs and allows friends to send each other haptic signals by connecting them with their phones to Bluetooth. When one user taps on their Hey Bracelet, the other one will squeeze and vibrate to let the wearer know their partner is sending them a hug. Similar capabilities are also possible with Apple Watches, where you can send a heartbeat to anyone via iMessage (Filipowicz & Keller, 2018). Another bracelet developed is CoupleVIBE (Bales, Li, & Griwsold, 2011), which sends a vibrotactile cue to a partner to signal that their long-distance significant other has arrived to or departed from a frequented location.

The HaptiHug is another kind of wearable for social touch (Tsetserukou, 2010); it is similar to a belt and features soft hands made from rubber-sponge material. The creators developed an animated version of a hug and integrated it into the online game Second Life. Real-world users can connect and interact in this virtual world. By wearing the HaptiHug while playing the game, users can feel squeezed and have pressure applied on their back when their virtual characters hug each other, thus making the experience more immersive.

Another kind of wearable technology for mediated social touch is the Hug Shirt (Genz, 2007), which has embedded sensors and actuators. The sender and the receiver each wear a Hug Shirt, and the system aims to create the sensation of physical contact between the wearers. The sensors capture the sender's contact location, strength, warmth, and heartbeat, and the actuators attempt to recreate these sensations on the receiver's shirt through heating, vibration, and inflation. The shirts send messages to each other via Bluetooth by connecting to the mobile phones of the users.

A final wearable we will discuss is the Huggy Pajama (Teh et al., 2008). These pajamas are meant to be a hugging communication system between children and parents. A parent hugs a doll embedded with sensors, and the child wearing the pajamas feels virtually hugged. The pajamas are actuated by air inflation with a compressor located outside of the pajamas. Unfortunately, this compressor is loud and can be disruptive to a child trying to sleep, though the provided compression was reported to be enjoyable.

2.2.0.2 *Comfort Objects*

Many people purchase "Friendship Lamps" because the idea of technology-mediated social touch appeals to users, even without research supporting their claimed benefits ("Friendship Lamps", 2020). These lamps work in pairs. When one user turns on theirs, the partner's lamp lights up. Touching either lamp causes both lamps to change colors to indicate to the partner that the other person is thinking of him/her. Another group of researchers took the idea of a tele-lamp one step further by giving it an anthropomorphic shape intended for affective interaction (Angelini et al., 2015). This lamp has a face, can change colors, and displays different emotions to reflect a distant loved one's emotional state. A user can change the lamp's emotional state by performing gestures on it, for example kissing the lamp. While couples can use this lamp for long-distance relationships, it can also be used alone as a single companion.

Similar to the friendship lamps, other researchers developed a plush pillow, the Macaron (Nunez et al., 2017), that uses infrared photo-reflective sensors to detect when a user is hugging it. It then sends a message over Bluetooth to the partner pillow, which lights up with blue-colored LEDs and blinks to indicate the intensity with which the partner hugged the other Macaron. Similarly, another group of researchers created The Hug (DiSalvo et al., 2003), a pillow whose shape is derived from a child wrapping its legs around an adult. The Hug also works in pairs. When one user hugs or strokes their Hug, the partner Hug will light up, vibrate, jingle, and heat up to indicate that someone is sending a hug.

2.3 ROBOTS DESIGNED FOR SOCIAL TOUCH

Social touch occurs in many contexts. Several researchers have devoted effort to enabling robots to shake hands with humans as naturally as humans do with each other, e.g., (Arns, Laliberte, & Gosselin, 2017; Wang, Peer, & Buss, 2009). Other researchers have worked to allow humans to connect with robots in a more light-hearted and playful manner through high-fives and hand-clapping games (Fitter & Kuchenbecker, 2016b, 2019; Fitter et al., 2020). All these interactions feature ways to help humans and robots interact with each other both socially and physically. This interest in enabling social-physical human-robot touch is not unique; many researchers have taken different approaches to this goal. Two of the more common approaches that are relevant to our research on HuggieBot involve social touch with robotic animals and human-robot hugs.

One such group of researchers expanded this research trend by creating the Haptic Creature (J. Chang, MacLean, & Yohanan, 2010; S. Yohanan & MacLean, 2012; S. J. Yohanan, 2012), which sits in the user's lap and looks like a large mouse. It uses a force-sensing resistor network embedded in its fiberglass shell and conductive external fur plus machine-learning techniques to detect and classify contacts. Users can perform gestures on this creature to convey different emotions, and the creature can mirror that emotion to the user. Users can feel it breathe and respond to their gestures, making it more interactive and life-like than previous robotic creature companions.

The PARO robot is another animal-based robotic platform used to provide social touch (W. L. Chang, Šabanović, & Huber, 2013; Šabanovic et al., 2013). This robot looks like a baby seal and is used as a socially assistive robot for therapy with dementia patients. It has dielectric tactile sensors under its fur that can detect force location and magnitude. It also has light, temperature, sound, and posture sensors. In response to a user, PARO can move its head and flippers as well as make sounds. PARO has been shown to increase older adults' activity levels and physical interaction with other humans well after its novelty has worn off.

In the Haptic Intelligence Department at the Max Planck Institute for Intelligent Systems, a group is covering a SoftBank Robotics NAO robot in custom fabric-based tactile sensors and a soft koala suit (Burns et al., 2021). This robot is meant to serve as a companion for children with autism, detecting where and how it is being contacted by the child and adapting its behavior accordingly.

Cooney et al. worked with a robot called Sponge Robot (Cooney et al., 2010). These researchers were interested in using internal inertial sensors, rather than external sensors, to recognize what full-body gestures humans were manipulating the robot to do. By observing humans interacting with Sponge Robot in free play, the researchers categorized the 13 most common interactions, which included hugging. Because of its small size, Sponge Robot, like Huggable, is able to be hugged but is unable to hug the user back. Later, Cooney et al. built on this work using both a hand-held robot, Elfoid, and a life-size robot, Kirin (Cooney, Nishio, & Ishiguro, 2012). This later work was focused on discovering typical human touches toward a humanoid robot. The 20 most common touches were categorized into "affectionate," "neutral," and "unaffectionate" categories. Hugging was found to be one of the most frequent affectionate touches humans wanted to express. Again, the robots in these studies did not actively hug the participant back; however, the fact hugging was a common interaction in both studies clearly demonstrates a desire among users to express affection to robots through hugging. These two studies support our belief that a robot that hugs users back would be well received.

In the autism community, there are many touchable systems for children with which to touch, hug, and play. One such system is Roball, which is a rolling robot ball whose external shell's rotation propels its motions (Laplante et al., 2005; Michaud & Caron, 2002). The robot's robust, safe, and simple design is created not only to appeal to and withstand children and provide for exciting play situations, but it also is particularly well suited for interacting with children with cognitive difficulties, like autism (Michaud et al., 2007). QueBall is a modified version of Roball that shares similar properties but allows for additional interaction modalities like touch sensing, sound, and colored lighting (Salter, Davey, & Michaud, 2014). The researchers developed five behaviors that address different autism-specific issues to encourage autistic children to play with QueBall. These behaviors helped autistic children work on learning, physical play, motor skills, and communication. Both Roball's and QueBall's success with children in the autism community reflects the importance of robustness and playfulness in designing a touchable system. We apply these learnings to our robot, ensuring that it is robust to withstand rough interplay situations. It is playful and fun so that users enjoy the hug interaction and find it enjoyable that they want to repeat the exchange and get another hug.

2.3.1 *Tactile Experience of Robotic Touch*

Some researchers focus on the how robots look and behave, but another important area of study is how robots feel. Shiomi, Nakagawa, et al. (2017) were interested in the connection between human-robot physical interaction and human effort. They gave participants a monotonous task to complete either with or without touch interactions from a robot. In cases with touch interactions, the robot held out its hand and asked the participant to hold it. In this case, the contact was initiated by the human, who then moved to hold the robot's hand. The researchers found that when contact occurred, the effort exerted by the human increased, both in terms of the number of actions taken and the length of time they worked. Thus, a hugging robot may be positively perceived and could be used in physical therapy or exercise to encourage human effort.

In contrast to the human-initiated touch described above, Chen et al.'s research featured robot-initiated touch in a clinical setting (T. L. Chen et al., 2014). While in some cases the robot gave a warning before touching the person, it differs from Shiomi et al.'s experiment because the robot reached to touch the participant, rather than having him/her contact the robot first. In this case, the robot touched a person's arm to either clean or comfort. Users preferred instrumental (cleaning) touch over affective (comforting) touch, which is the same result found between patients and human nurses. The perceived intent of the robot greatly affects the participant's acceptance of the touch. In robot-initiated touch, the authors suggest that any warning prior to contact should be carefully worded. It appears human-initiated contact with a robot is more readily accepted. Taking this preference into account, our experiment features human-initiated contact. The robot first asks the participant for a hug, then the participant must walk to the robot for the hug.

While Chen et al. found that human-initiated touch is preferred in a clinical setting, Cramer et al. were interested in what touch is preferred in a more relaxed setting (Cramer et al., 2009). Cramer et al.'s research featured three robot-initiated contacts, one human-initiated contact, and a no contact situation. These researchers found that users preferred when the touch was robot-initiated rather than human-initiated. Users who already had a positive attitude towards robots in general perceived a closer personal connection with the robot. These participants also found robots that interact by touch to be more natural than robots that don't.

The tactile experience of the user is not limited to who or what initiates contact; the physical qualities and visual appearance of that contact also matter. Williams and Bargh's coffee temperature result was confirmed in robotics when Park and Lee altered the skin temperature of a dinosaur robot (E. Park & Lee, 2014). They found user perceptions of the sociability of the robot increased with warmth. Nie et al. found similar results when participants watching a horror movie either held hands with a warm robot, held hands with a cold robot, or did not hold hands at all (Nie et al., 2012). Participants who experienced physical warmth viewed the robot with increased friendship and trust. Furthermore, using a very mechanical looking robot (RoboSapien) to perform a human-like interaction yielded negative reactions from participants in this experiment. These results led us to cover our mechanical looking robot with fuzzy clothes. The user acceptance and enjoyment benefits of warmth in both Park and Lee's and Nee et al.'s studies led us to examine the impact of robot warmth in our study of robot hugging

2.3.2 *Robot Touch for Social Enjoyment*

A common way humans warm up to each other is by interacting physically for social enjoyment, through activities such as high-fives, games, and dancing. Several researchers have focused on the social aspects of physical human-robot interactions by attempting to recreate these common forms of expression. For example, Romano and Kuchenbecker's demonstration with the PR2 included a range of social-physical interactions like high fives, fist bumps, and hugs (Romano & Kuchenbecker, 2011). Their robot hug was used as a starting point for this research. Fitter and Kuchenbecker studied human-human play in order to program the Baxter robot to play hand-clapping games with humans (Fitter, Hawkes, & Kuchenbecker, 2016). Both these projects attempt to recreate a fun human-human interaction with a robot.

Partner dancing is a close social-physical interaction that researchers have recreated by replacing one human with a robot. Peng et al. discuss the ways that several groups have looked into cooperative human-robot dance (Peng et al., 2015). Pattern Ballroom Dance Robot (PBDR) and its prototype, Ms-DanceR, are two waltzing robots created to bring human-robot collaboration closer than ever (Kosuge et al., 2008). These robots have wheels, rather than legs, to avoid the problem of balancing, as well as a small front base, to allow the human to get close enough to the robot for the required dance embrace. The SpiderCrab robot is a robotic arm that uses improvisational dance to interact with its human partners (Wallis, Popat, & McKinney, 2010). Dancers feel that because the robot responds to their spontaneous movements, it behaves as a real human dance partner might. A similar responsive element is integrated into our system, with the human being able to control the duration of the hug. Dance and hugs are two different forms of non-verbal social interactions. The various different approaches show there is a desire for more social interactions between humans and robots.

3

RELATED WORK

Hugging: the truest form of giving and receiving
— Carol 'CC' Miller, Positive Impact Coach

In recent years, there has been a dramatic increase in the number of researchers developing and studying hugging robots. High press coverage shows this topic is of great interest to the general public. Interestingly, researchers are taking many different approaches to create robotic systems that can receive and/or give hugs.

In this chapter, we provide a brief review to familiarize the reader with the other hugging robotic solutions, which have greatly inspired this thesis. Such robots were synthesized into three main categories: zoomorphic hugging robots, small form-factor hugging robots, and human-sized anthropomorphic hugging robots. At the end, we also touch on the only other study to investigate the physiological effects of non-human hugging.

3.1 ZOOMORPHIC HUGGING ROBOTS

Zoomorphic hugging robots are fun, and they can particularly appeal to children. Thus far they have been tele-operated rather than fully autonomous, and they sit on the floor, which require users to awkwardly crawl to them to receive a hug.

Shiomi et al. ran two Wizard-of-Oz-style experiments with a large teddy bear robot that was covered with polypropylene cotton and had two elbows powered by easily backdrivable motors (for participant safety) (Shiomi, Nakata, et al., 2017a, 2017b). The robot always sits on the floor and introduces itself. One experiment varied whether the robot asked the person for a hug after this introduction. The rest of the interaction was a conversation wherein the robot asked the participant to tell stories about him or herself. Significantly longer interaction times and significantly more personal self-disclosure occurred when participants hugged the robot (Shiomi, Nakata, et al., 2017b).

In Shiomi et al.'s second experiment, the robot requested a hug from all participants, but it reciprocated the hug (squeezed them back) in only some instances. At the end of the subsequent conversation, the robot asked whether the participant would like to donate to earthquake victims. Shiomi, Nakata, et al. (2017a) found that users whose hugs were reciprocated tended to donate more money than those whose hugs were not, although this trend did not reach statistical significance. Participant acceptance of this robot hug behavior encouraged us to use the design as a model for our own experiment. While the system currently requires a human operator, the hug itself is meant to come from the robot, not another user; thus, we consider it a human-robot hug rather than a technology-mediated hug. This floor-sitting robot requires users to crawl toward the robot to hug it.

HugBot is a large panda bear stuffed animal (Hedayati et al., 2019) similar to Shiomi et al.'s life-size teddy bear robot. Hugbot also sits on the floor and requires adults to crawl to it to receive a hug; children can simply bend down. The robot has pressure sensors (four on the chest and two on each arm) to record how much pressure is exerted and an inner t-shaped wooden structure upon which the two soft robotic arms are attached. The sensors and actuators are situated inside a large stuffed animal. In their study, these researchers found that zoomorphic robots kept children more engaged compared to non-zoomorphic robots.

3.2 SMALL FORM-FACTOR HUGGING ROBOTS

Smaller hugging robots have been created to provide comfort, but they typically cannot actively hug the user back. Their small size makes these huggable systems inherently safer than larger devices, but it also prevents them from providing the benefits of social touch to the user because they cannot administer deep touch pressure therapy (Edelson et al., 1999).

DiSalvo et al.'s the Hug is a plush comfort object that works in pairs to provide users "tele-hugs" (DiSalvo et al., 2003). Its shape mimics a child wrapping their limbs around an adult. The Hug plays a melody and its stomach glows to alert the user that its partner is sending a hug. While one user strokes their Hug, the Hug with the other user vibrates to match. Throughout the course of voice conservations held with the Hug, it will warm up to a comfortable heat. The Hug lacks the ability to wrap its arms tighter around the human, and it requires a human partner for the user to feel emotional support.

Stiehl et al.'s teddy bear robot, the Huggable, employs a more interactive design, and is meant to accompnay children with long stays in the hospital (Stiehl et al., 2006; Stiehl et al., 2005). Huggable's temperature, force, and electric field sensors are concealed under a layer of silicone and fur for an enjoyable tactile interface. The robot is small enough to be held in a child's arms, it can detect where and how it is being touched, and it can move its head and arms in response. With cameras, microphones, and a speaker, it records and engages the person using it, while providing helpful information to a remote caregiver. Huggable has sensitive comfortable skin, but it still falls short of accurately replicating a human hug due to its miniature size and the fact that it cannot measure and reciprocate the pressure with which it is being hugged.

Another group of researchers developed Hugvie (Nakanishi, Sumioka, & Ishiguro, 2020; Sumioka et al., 2013; Yamazaki et al., 2016), a small pillow designed to have the approximate shape of a person. The pillow's head contains a small pocket in which a user can place a cell phone. Users then hug the pillow while talking on the phone to their far-away loved one. Researchers found that using this technology can help reduce the hugger's anxiety. They also found the interpersonal touch from using Hugvie can help improve the hugger's impression of a third person based on hearsay information given by the remote partner. As the authors categorize the Hugvie as a "robot," we have chosen to compare it to other small-form factor hugging robots; however, the Hugvie does not have any actuation. It cannot move, record, or respond to anything the user does. It cannot be programmed to perform any actions or make any sound.

3.3 HUMAN-SIZED ANTHROPOMORPHIC HUGGING ROBOTS

Human-sized anthropomorphic hugging robots take many different shapes and sizes.

Miyashita and Ishiguro used a wheeled inverted pendulum humanoid robot named Robovie-III that hugs in a three-step process while maintaining its balance (Miyashita & Ishiguro, 2004). First, the robot opens its arms. The robot waits for the user to approach and measures the distance between itself and the human by ultrasonic range sensors. When the distance is appropriate, the robot wraps its arms around the person. Finally, the robot opens its arms again. During these three steps, the robot is maintaining its balance with a balancing controller. This research does not address how the robot determines when to switch between the second and third steps, thereby deciding the duration of the hug. It also appears the robot uses the human to balance itself, which is potentially uncomfortable for humans. It is made of metal without any soft, cushioned material, so the tactile experience may not be enjoyable for the human.

Yamane et al. have been issued a patent (US 9802314) for Disney Enterprises, Inc. to create their own version of a huggable robot (Yamane, Kim, & Alspach, 2017). This robot is designed for human interaction, presumably within theme parks. It features a rigid inner structure with specific elements made of softer material to create a deformable exterior in areas that would contact a human. This robot attempts to match the pressure an external user applies using pressure sensors. The wording of the patent is ambiguous as to whether this robot will be autonomous or tele-operated, but it seems more likely intended to be tele-operated. However, as the users expect the hug to come from the robot itself, not another person, we identify it as human-robot interaction, rather than technology-mediated hugs. The physical appearance of this robot matches that of the character Baymax from the Disney movie "Big Hero 6" (2014). This patent application supports the belief that there is great interest in furthering human-robot interaction to include more pleasant physical exchanges, particularly human-robot hugging.

Kaplish and Yamane (2019) created a teleoperated humanoid robot for physical human-robot interaction. Their robot uses two Franka Emika robotic arms, covered by 3D printed shells. These shells are then covered in polyurethane foam for a soft and enjoyable tactile experience for the user, and bellows cover the gaps in the elbows and neck. The shells also have embedded optics-based force sensors, which the authors use to estimate the hug tightness. The upper body is then covered in custom stretchable fabric. In this study, a human operator wears a suit covered in eight IMUs and the same force sensors that are found on the robot, as he hugs a mannequin. The robot then hugs a dummy while being controlled by the pre-recorded sensor measurements on the human operator's suits. In this experiment, the robot never hugged another person and did not hug in real time. The main focus was testing the sensor mapping capabilities from the human operator's suit to the robot's force and motion control.

Later, Campbell and Yamane (2020) build on Kaplish and Yamane (2019)'s work using the same robotic platform, this time testing with users directly hugging the robot. The robot was equipped with 61 force sensors. They combine temporal and spatial sparsity with a learning from demonstration framework for whole-body haptic interactions. They trained a model with teleoperated data collected from 121 sample hugs from four participants. They found that it can generalize well to unseen hug styles and the six new human hugging partners. The authors created solutions to edge cases when the human or robot hugging partner did not perform as expected. Such edge cases included the robot hugging the air, a delay before closing the arms for a hug, or hugging without making contact ("air hug"). This anticipation of and preparation for users and the robot performing unexpectedly inspired our decision to integrate edge cases for our own hugging robot in Chapter 7.

For my master's thesis, we added padding, heating pads, and a soft tactile sensor to a Willow Garage Personal Robot 2 (PR2) to enable it to exchange hugs with human users (Block, 2017) we manually adjusted the robot to match the height and size of each user, a process that takes time and prevents spontaneous hugging. We also initiated every hug for the user. The PR2 was successful in adapting to the user's desired hug duration through the use of the stretchable tactile sensor (Y. Chen et al., 2016), but the user had to place their hand in a specific location on the PR2's back, which was not natural for all users, and the user had to press this sensor to tell the robot to release them. We tested combinations of soft and warm hugs and tested the pressure the robot should apply and how long the hug should last. Some users also criticized the size and shape of this robot as being awkward to hug. Further details, results, and analyses on this study can be found in Chapter 4, as much of this work was conducted during my doctorate.

3.4 EVALUATION OF THE EFFECTS OF ROBOT HUGS

Eckstein et al. (2020) review the current literature on touch performed by different agents (human, animal, and robot) on a person. They find that the effect is calming and stress-relieving regardless of the agent performing the touch. This conclusion supports our hypothesis that a hug from a robot will have a calming effect similar to a human hug.

The closest researchers have come to proving that robot hugs affect the human body similarly to human hugs is the work of Sumioka et al. (2013), who created Hugvie (mentioned in Section 3.2). These researchers collected initial blood and saliva samples upon participant arrival. After a five-minute rest period, each user had a fifteen-minute phone conversation with a stranger. One group hugged a pillow with an embedded cell phone throughout the conversation, while the other group did not. Following the fifteen-minute conversation, blood and saliva samples were collected again. The group that hugged the pillow during the conversation had a greater decrease in cortisol (stress hormone) levels, similar to the effects of self-soothing. To the best of our knowledge, there is no published work on how an interactive robot hug affects human physiology.

4

HUGGIEBOT 1.0: SOFTNESS, WARMTH, AND RESPONSIVENESS IMPROVE ROBOT HUGS

Hi, I'm Olaf, and I like warm hugs!.
— Olaf, the snowman from Disney's Frozen (2013)

4.1 INTRODUCTION

This research sought to determine an effective way in which a robot should hug a human, which is a delicate social-physical interaction (Fig. 1.1). This chapter builds on previous work in the field of social human-robot interaction (HRI), as discussed in Chapter 3. Section 4.2 outlines the four hypotheses tested. A description of the robot outfits created, sensors used, and robot programs developed can be read in Section 4.3. Details of the experiment setup and participant breakdown can be found in Section 4.4. Results from the experiment are presented in Section 4.5 and discussed in Section 4.6. We then summarize our conclusions of this work in Section 4.7.

4.2 RESEARCH QUESTION

This research question builds on the aforementioned literature and matches common human social conventions to discover which factors can create a pleasant robot hug. A pleasant robot hug is defined as a hug that is comfortable and enjoyable to the individual such that when there is no longer a requirement to hug, the user chooses to go back and hug the robot again, thereby showing interest and the belief that it is worth their time. Robot hug pleasantness is evaluated through self-report survey responses and user behavior.

This chapter tests the research question that ***a soft, warm, touch-sensitive humanoid robot can provide humans with satisfying hugs by matching their hugging pressure and duration.*** The first two investigated factors pertain to physical aspects of the robot's body, and the second two relate to how it moves during the hug.

The first factor is softness. Whatever object is hugged should be enjoyable to hug, by deforming somewhat, perhaps similar to human tissue. For this experiment, a layer of foam and fluffy fabric covers the robot, similar to DiSalvo et al.'s the Hug (DiSalvo et al., 2003). Another part of what makes a hug enjoyable is the warmth from another human body. We thus heat the exterior of the robot by adding heating elements so it is easily recognized as warmer than ambient temperature.

Next, for a robot to give a pleasant hug, we believe it should reciprocate the amount of pressure the human applies (up to a certain threshold for safety). We calibrate the robot to the size of each user, then vary how tightly it hugs the person. The final factor to a pleasant hug is that it lasts an appropriate duration, which may vary for different people. We use a tactile sensor to measure when contact is made and broken, and we terminate the hug before, at, or after the user releases the robot.

We break down this overall research question into four sub-statements to better test their validity:

H1. Participants will prefer hugging a cold, soft robot rather than a cold, hard robot.

H2. Users will prefer hugging a warm, soft robot rather than a cold, soft robot.

H3. Participants will prefer a robot that hugs with a medium amount of pressure, rather than hugging too loosely or too tightly.

H4. Users will prefer a robot that releases them from a hug immediately when they indicate they are ready for the hug to be over, rather than the robot releasing the hug before or after this time.

4.3 SYSTEM DESIGN AND ENGINEERING

This robot implementation uses a Willow Garage Personal Robot 2 (PR2) to exchange hugs with human users. As seen in Fig. 4.1, the PR2 has two seven-degree-of-freedom arms and a head mounted to a torso that can move up and down. Although the robot has a large mobile base, it was kept stationary throughout this study to focus on the main hugging interaction, which is delivered with the arms. The PR2 has a hard metal exterior, with cloth covering some of the arm surfaces.

Three different physical conditions were created, as shown in Fig. 4.2: a Hard-Cold, a Soft-Cold, and a Soft-Warm robot.

- **Hard-Cold:** The robot does not have any additional padding layers.

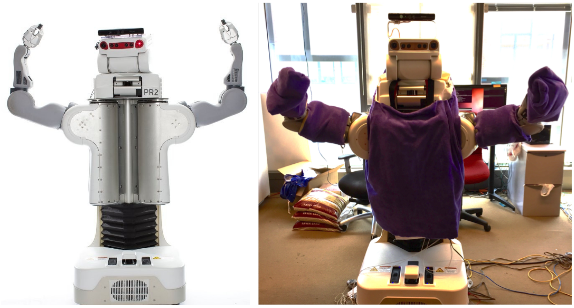


FIGURE 4.1: The left image shows the PR2 in the Hard-Cold condition, and the right image shows the robot in the Soft-Warm condition, wearing the custom-made outfit.

- **Soft-Cold:** The robot wears layers of foam, cotton, and purple fluffy polyester.
- **Soft-Warm:** This condition is the same as the Soft-Cold condition with the addition of heat. We made a separate, identical purple fluffy polyester layer. The cords were removed from a Sunbeam Quilted Fleece Heated Blanket (model number BSF9GQS-R727-13A00, 218.44 cm. × 228.60 cm. × 1.27 cm.). These cords were then sewn into channels between cotton layers to heat the chest and back of the robot. A Thermophore MaxHeat Deep-Heat heating pad (model number: 155, 68.58 cm. × 35.56 cm. × 1.27 cm.) was placed on top of the cotton layer and below the polyester layer on the chest of the robot for added warmth. The final warming components were four chemical warming packs, (model number: TT240-AMZ, 6.98 cm. × 8.64 cm. × 1.27 cm.) which were placed on both upper arms and forearms of the robot, on top of the foam and beneath the polyester layer.

A Hard-Warm robot was not tested because the heating elements somewhat soften the robot and divulge the thermal experimental variable to users. The robot outfit was designed to be gender neutral to be more universally accepted and to limit the number of experimental variables.

4.3.1 *Tactile Sensors*

In order to tell when users made and broke contact with the robot, we needed a haptic sensor. This chapter uses a stretchable tactile sensor made

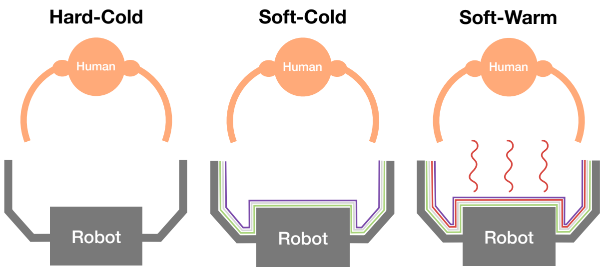


FIGURE 4.2: An overhead view of the three experimental conditions to be tested.

by Chen et al. (Y. Chen et al., 2016). This team created a strain-sensing skin by spray-coating latex with an exfoliated graphite piezoresistive sensing paint. To determine the ideal placement location, initial sensor trials were conducted with the sensor mounted on the robot's chest foam. The sensors were placed in series with a 10 kΩ resistor to act as a voltage divider.

For the final experiment, the tactile sensor was moved from the chest to the back for several reasons. First, the base of the PR2 is very large, so a person had to lean very far towards the robot to make enough contact with the chest-mounted tactile sensor, which some pilot participants were hesitant to do. Next, the PR2's arms make up the chest area, which rotate as the arms moves. This caused the foam covering the chest area to shift as well. With the foam and sensor pulled between the two arms, it became difficult for a user to make contact with the sensor. Finally, if the robot was hugging in the "too tight" condition and the participant was unable to break contact with the chest sensor, the robot would have no way of knowing when the human wanted to be released from the hug. For these reasons, it was determined that the best location for the sensor would be on the upper back of the robot, where the users' hands would naturally be placed. In this manner, the sensors are able to detect touch through strain that is induced by contact, and had an even stronger signal than the original placement.

To smooth the data, we applied an infinite impulse response low-pass filter. The equation for this filter is as follows:

$$V_{\text{smooth,k}} = w \cdot V_{\text{k}} + (1 - w) \cdot V_{\text{smooth,k}-1} \tag{4.1}$$

where $V_{\text{smooth,k}}$ is the current smoothed voltage sample, w is the filter weight, for which we used 0.08, giving a filter bandwidth of 0.70 Hz, V_{k} is

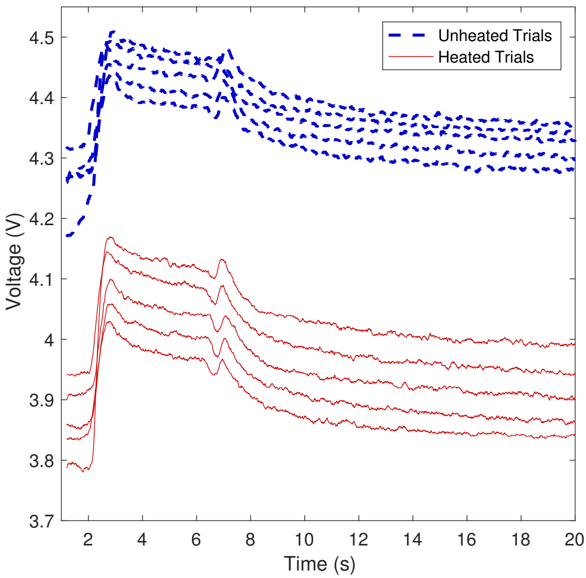


FIGURE 4.3: The filtered data from the unheated and heated tactile sensor trials, with the sensor mounted on the chest.

the current raw voltage sample, and $V_{\mathrm{smooth,k-1}}$ is the previous smoothed voltage sample.

Fig. 4.3 shows the filtered data from the sensor trials. It is important to note that temperature greatly affected the resistance of the sensor, as can be seen by the five unheated trials having higher resistance than the five heated trials in Fig. 4.3. While the magnitude of the heated and unheated trials varied greatly, the overall slope is similar.

We take the derivative of the voltage over time according to the following equation:

$$\frac{dV_k}{dt} = (V_{\mathrm{smooth,k+1}} - V_{\mathrm{smooth,k}})/(t_{\mathrm{k+1}} - t_{\mathrm{k}}) \tag{4.2}$$

where $\frac{dV_k}{dt}$ is the current value of the derivative of the voltage over time, $V_{\mathrm{smooth,k+1}}$ is the following smoothed voltage sample, and $V_{\mathrm{smooth,k}}$ is the current smoothed voltage sample. $t_{\mathrm{k+1}}$ is the following time point, and t_{k} is the current time point.

By taking the derivative of the voltage over time, all the trials collapse into each other. We can then clearly identify two defined peaks which indicate the initiation and conclusion of a hug.

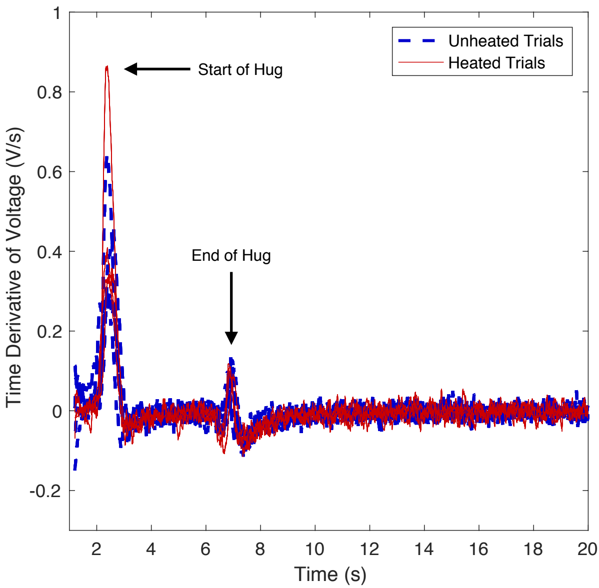


FIGURE 4.4: The derivative of the voltage over time for the tactile sensor trial data, from Fig. 4.3.

Fig. 4.4 shows that the tactile sensor measurements can clearly identify and inform the robot when human contact is made and broken. Using this information, the robot releases the embrace based on the pre-programmed duration with which it is set to hug.

4.3.2 *Robot Program*

Starting with the aforementioned demonstration developed by Romano and Kuchenbecker (Romano & Kuchenbecker, 2011), we edited the hug code in several ways to create a more natural hug, by adjusting the tightness and duration of the hug. An Arduino samples and communicates the tactile sensor data to ROS over serial at 173 Hz, which is then filtered and read into the hug code. Ten different hugs (nine for the behavioral changes, and one used for all three physical changes) were created.

At the start of the experiment, mannequin mode was used to find the joint angles for a comfortable hug based on the participant's body size. The arms were manually wrapped around each user. These unique joint angles were collected and used for the "just right" hug for each individual.

Decreasing the first angle (shoulder pan) by 0.05 radians, and the fourth angle (elbow flex) by 0.2 radians created the "too tight" hug for the right arm. Increasing these two angles by the same amount created the "too tight" hug for the left arm. The "too loose" hug joint angles were set loose enough so they did not touch any participants; thus they were left standard for all trials, and can be seen below, recorded in radians and rounded to two significant figures. These changes in the shoulder and elbow angles to create the "too tight" and "too loose" conditions were experimentally determined.

θ_r= [-0.19, -0.20, -1.25, -1.58, 0.02, -0.49, 0.11]

θ_l = [0.17, 0.05, 1.22, -1.35, 1.22, -1.34, -0.83]

Even though the robot's arms did not touch any participant, the "too loose" condition can still be considered a hug for several reasons. First, during all hugs, the participant's arms fully encircled and touched the robot. Participants always touched the robot to ensure they would be able to activate the robot's tactile sensor to let it know when to terminate the hug. Therefore, this interaction satisfies the belief that a hug implies touch. Next, the robot's arms moved through the hugging motion and did fully encircle the participant, even though the arms didn't necessarily touch them. The robot was essentially giving the participant an "air hug" in this condition, which is common for people to give each other (for example when a person is sick and you want to comfort them but not get too close or touch so you don't also get sick).

To begin the hug sequence, we added that the robot lifts its arms over the course of four seconds and asks the participant for a hug by saying "Can I have a hug, pleeeease?" After waiting two seconds, the robot then closes its arms for five seconds, using the ROS joint trajectory action, to hug the person according to the pre-assigned condition of pressure to apply. For a too short hug, the robot waits one second from when it closes before it releases. Otherwise, the robot continually monitors the derivative of the voltage values until it notices the second spike, with a threshold that was tuned during pilot testing to be 0.2 V/s. For a immediate release hug, it releases immediately when it notices the spike. For a too long hug, the robot releases five seconds after it notices the person releases. After the robot returns to the joint angles associated with the outstretched arms position in which it began (which takes five seconds), the robot drops its arms (which takes four seconds) to mimic a human motion.

The success of the experiment relies on the automatic detection of the start and the end of the hug, which is a new component of the system. To

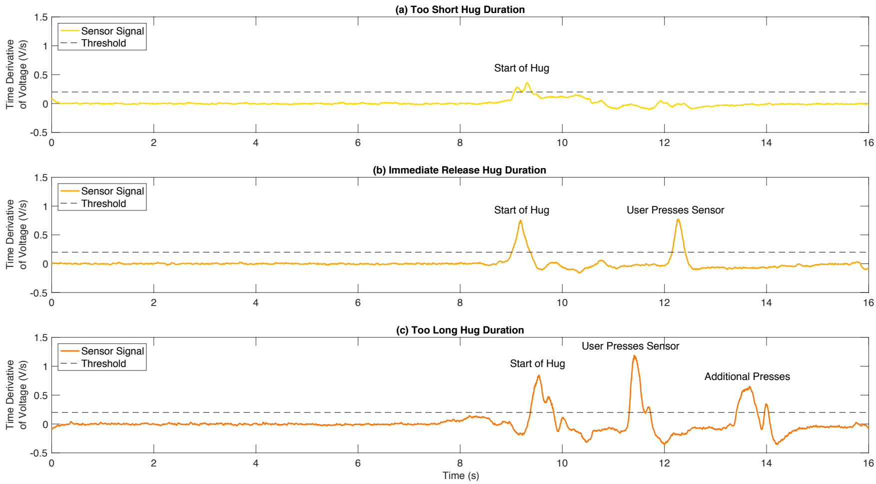


FIGURE 4.5: Tactile sensor data for three randomly selected trials from the experiment. Plot (a) shows a "too short" hug duration trial, plot (b) shows an "immediate release" hug duration trial, and plot (c) shows a "too long" hug duration trial, in which the user worried because the robot wasn't letting go and pressed on the sensor additional times.

verify that this method was reliable, the experimenter monitored the sensor data for every participant for every trial (excluding the Hard-Cold condition, which had no sensor data). No malfunctions were detected. Fig. 4.5 shows the time derivative of the tactile sensor measurements collected during three randomly selected trials to show that automatic detection of the start and end of the hug were reliable. The top plot shows the "too short" hug duration with only one spike at the start of the hug because the robot released before the user was ready to be released. The second plot shows the "immediate release" hug duration with two spikes indicating the start and end of the hug. The bottom plot shows the "too long" hug duration in which the user worried because the robot wasn't letting go and pressed additional times.

4.4 USER STUDY METHODS

All methods for this study were approved by the University of Pennsylvania Institutional Review Board under protocol # 827123. To recruit participants, the investigator circulated a recruitment email through mailing lists to which our lab has access. Potential users read the advertisement and contacted the investigator via email or phone to learn more and possibly schedule an appointment.

4.4.1 *Procedures*

A flowchart of the experimental procedure can be found in Fig. 4.6. After the potential participant arrived at the experiment site, the investigator explained the experiment, using the informed consent form as a guide. The potential user took several minutes to read over the consent form and ask questions. If they still wanted to volunteer, the user signed the informed consent form. Prior to continuing, the investigator verbally verified that the participant had full use of their arms and legs, had no uncorrected vision or hearing impairments, was fluent in English, and was a legal adult. After the user signed the consent form, the investigator turned on a video camera and began to record the experiment.

Next, the investigator introduced the robot as the personality "HuggieBot" and explained the key features of the Willow Garage PR2, such as the two emergency stops, one located on the back of the robot, the other on a remote control. The investigator introduced the participant to the user and explained how the trials work and how the user should be prepared to move. The participant was taught the involved hug motions in this step. The PR2 was then customized to the size of the person, while wearing its Soft-Cold outfit. By having the participant stand in front of the robot, the experimenter raised or lowered the body of the PR2 until the user said the robot's eye cameras were at their eye level. Then, the robot was put in mannequin mode and the arms were manually placed in a comfortable hug around the user. The joint angles for this specific motion were collected.

Finally, the participant was reminded of the possibility to terminate the experiment or the robot's movement at any time by verbal request. The investigator then prepared the robot for the first of the three physical attribute trials. While the experimenter was making the necessary changes, the participant was instructed to complete an opening survey based on their initial impressions of the robot. The questionnaire asked the participant to rate how much they agreed with each of the statements listed in Table 4.1 on a sliding scale from 0 (not at all) to 100 (a great deal). Note that these questions were asked before the user had experienced any active hugs with the robot.

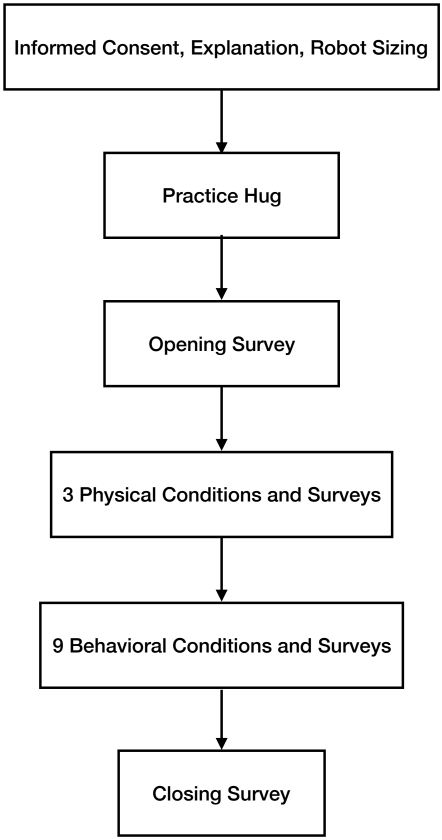


FIGURE 4.6: A flowchart of the experimental procedure.

Question
I feel understood by the robot
I trust the robot
Robots would be nice to hug
I like the presence of the robot
I think using the robot is a good idea
I am afraid to break something while using the robot
People would be impressed if I had such a robot
I could cooperate with the robot
I think the robot is easy to use
I could do activities with this robot
I feel threatened by the robot
This robot would be useful for me
This robot could help me
This robot could support me
I consider this robot to be a social agent

TABLE 4.1: The fifteen survey questions asked in the opening and closing questionaires.

Once all changes were completed, the participant was called out from behind the privacy screen and completed a practice run of the human-robot interaction involved in the experiment. During this practice hug, the robot was in the condition of the user's first trial. The goal of this practice was to ensure correct arm placement to activate the tactile sensor, acclimate the participant to the timing of the hug (from when the robot asked, "Can I have a hug, pleeease?" to when it closed its arms), and to address the worries users might have about hugging a robot. Some participants were initially hesitant to hug the robot and missed the first practice hug, so they generally appreciated having multiple chances to get accustomed to it before the start of the experiment. On average, people did 2-3 practice hugs before they felt comfortable beginning the experiment. The twelve tested conditions each participant experienced are listed in Table 4.2.

A within-subjects study was selected for this experiment for several reasons. First, we were most interested in the differences between the

Physical Conditions
Hard-Cold
Soft-Cold
Soft-Warm
Behavioral Conditions
Too Loose, Too Short
Too Loose, Immediate Release Duration
Too Loose, Too Long
Just Right Tightness, Too Short
Just Right Tightness, Immediate Release Duration
Just Right Tightness, Too Long
Too Tight, Too Short
Too Tight, Immediate Release Duration
Too Tight, Too Long

TABLE 4.2: A description of the twelve tested hug conditions, broken into the two parts of the experiment.

conditions rather than the overall response levels to a robot hug. We also preferred this design for its higher statistical power and the lower number of participants it requires compared to a between-subjects study (de Winter & Dodou, 2017). An image of a participant hugging the robot in the Soft-Warm condition can be seen in Fig. 4.7. A sample compilation video of participants hugging the robot during the experiment in various hug conditions is provided as supplementary material.

The three physical trials for the first part of the experiment consisted of the robot in either the Soft-Warm, Soft-Cold, or Hard-Cold condition. In all of these trials, the robot hugged with the "just right" pressure and hugged for a preset duration of three seconds before releasing. After each of these hugs, the user completed a brief quantitative survey concerning their perception of the robot attributes; these survey questions can be found in Table 4.3. Changing the robot outfit between trials took approximately five minutes. The participant sat at a desk behind a privacy screen and answered the survey questions during this time. Afterwards, the participant read the school newspaper until they were called for the next trial.

FIGURE 4.7: A participant hugging the PR2 with its custom Soft-Warm outfit during the experiment.

The nine behavioral trials for the second phase of the experiment all occurred with the PR2 in the Soft-Warm condition, but with it interacting with the participant in a different way. Here, the robot varied how hard (too loose, just right, too tight) and how long (too short, immediate release, and too long) it hugged the person. Prior to starting the second part of the experiment, participants were taught to press down quickly on the tactile sensor when they wanted to be released from the hug, similarly to how they might pat someone on the back at the end of a hug. Each trial again was followed by the same robot attribute questionnaire (Table 4.3).

Prior to any experimentation, the investigator developed a randomized order of trials for up to 40 participants, thereby eliminating presentation order effects. The two parts of the experiment were handled separately for randomizing the hug orders. Participants always experienced the three physical conditions first and the nine behavioral conditions afterwards. Therefore, the presentation orders of the hugs were randomized with respect to their category.

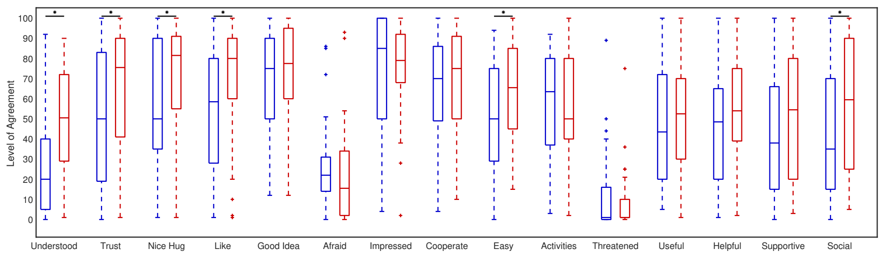


FIGURE 4.8: A comparison of the responses to the opening (blue) and closing (red) surveys. The top and bottom of the box represent the 25th and 75th percentile responses, respectively, while the line in the center of the box represents the median. The lines extending past the boxes show the farthest data point not considered outliers. The + marks indicate outliers. The black lines with stars at the top of the graph indicate where a statistically significant difference was found between the opening and closing survey.

At the end of the experiment, the participant completed the brief survey concerning the robot from Table 4.1, and they answered the four additional open-ended questions that can be seen in Table 4.4. These subjective measures were used for evaluation because we were interested in user preference, attitude, and opinion.

All slider-type questions in the surveys were based on previous surveys in HRI research and typical Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaires (Heerink et al., 2009; Weiss et al., 2008), and the free-response questions were designed to give the investigators any other information the participant would like to share about the experiment experience. At the end of the entire experiment, the user then completed a brief demographics questionnaire and was thanked for participating in the study. The investigator answered any questions the participant had and escorted the participant out of the experiment venue.

4.4.2 *Participants*

All participants were volunteers recruited from the Penn Engineering community from emails sent out through listservs. We ran two users as pilot participants to refine the experimental methods; their data are excluded from analysis because they were not given the same instructions as the later participants. A total of 30 people participated in the study: 14 male, 15

Question
Which hug are you evaluating?
This robot behavior seemed (unsafe – safe)
I think the robot is (anti-social – social)
I think the robot is (selfish – caring)
Hugging the robot makes me (unhappy – happy)
Hugging the robot makes me feel (uncomfortable – comforted)

TABLE 4.3: The survey each participant completed after each hug.

Question
What aspects of this activity did you enjoy?
What aspects of this activity were most challenging?
Why would or wouldn't you want to do this activity with a robot?
What other activities would you want to do with this robot?

TABLE 4.4: The open-ended questions each participant answered at the end of the closing questionaire.

female, and 1 who identifies as "other". Ages of participants ranged from 21 to 54 (M = 26.8, SD = 8.5). 23 participants were engineering students, and 7 were faculty or staff associated with the university. The majority of participants (25) had a technical education/background, while 5 did not. Users also had varied amounts of experience with robots. When asked on a sliding scale from 0 = "novice" to 100 = "expert," many had some experience with robots in general (M = 52.5, SD = 31.2), while fewer had experience with the PR2 specifically (M = 15.8, SD = 26.5).

4.5 USER STUDY RESULTS

We analyzed five main sources of data to understand how robots should hug: the opening and closing experiment survey, the post-trial survey from the first three trials, the post-trial survey from the last nine trials, verbal comments made by participants during the experiment, and the free-response comments written by users at the end of the study.

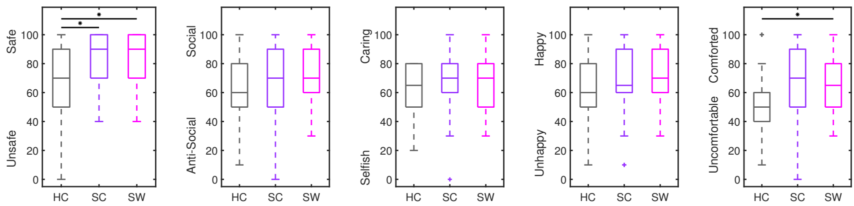
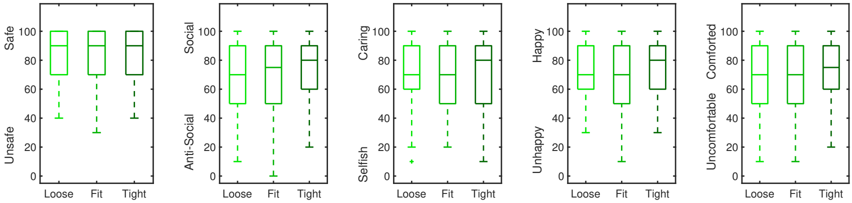


FIGURE 4.9: A comparison of the responses to the survey questions after the first three hugs, changing physical conditions. The grey represents the Hard-Cold condition (HC), the purple color represents the Soft-Cold condition (SC), and the pink color represents the Soft-Warm condition (SW). The top and bottom of the box represent the 25th and 75th percentile responses, respectively, while the line in the center of the box represents the median. The lines extending past the boxes show the farthest data point not considered outliers. The + marks indicate outliers.The black lines with stars at the top of the graph indicate statistically significant differences.

4.5.1 *Opening and Closing Survey Results*

As mentioned earlier, the participant was presented with a set of 15 questions prior to any experimentation. This opening survey was answered by the participant based on initial impressions of the robot after a brief introduction. Following the conclusion of the entire experiment, the user was asked the same set of 15 questions, which can be found in Table 4.1. Box plots of the responses to the opening and closing survey questions are shown in Fig. 4.8. For all analyses we used $\alpha = 0.05$ to determine significance. Using a paired t-test comparison of the opening and closing survey, we found that users felt understood by ($p < 0.005$), trusted ($p < 0.005$), and liked the presence of the robot ($p < 0.005$) significantly more after the experiment. People also found robots to be nicer to hug ($p < 0.005$), easier to use ($p = 0.0087$), and more of a social agent ($p < 0.005$) than they initially anticipated.

Users' opinions about whether using the robot was a good idea did not change significantly over the course of the study, nor did their fear of breaking the robot. They also did not feel that people would be significantly more or less impressed if they had a hugging robot after concluding the experiment. Participant beliefs about their ability to cooperate with the robot were not significantly affected by the experiment. The course of the study did not significantly change user opinions about whether they could



do activities with the robot, how threatened they felt by the robot, how useful, how helpful, or how supportive they thought the robot would be.

4.5.2 *Survey Results from Physical Trials*

The first three hugs showcased changes in the physical properties of the robot. Box plots of the responses to the five questions participants were asked after each hug can be seen in Fig. 4.9. The ratings after each trial were analyzed in MATLAB 2017a using a one-way repeated measures analysis of variance (`ranova`). Afterwards, we ran a Tukey posthoc multiple comparison test to determine which conditions were significantly different from each other (`multcompare`). Our data satisfies all the assumptions of a one-way repeated measures ANOVA.

A significant difference in the perceived safety of the robot was noticed when the robot was covered in foam (both Soft-Cold and Soft-Warm), compared to when it was not ($F(2,58) = 5.28$, $p = 0.0078$), with both softer conditions being preferred. There was not a significant difference noticed between the two soft conditions. The statistical significance between the Hard-Cold and Soft-Cold conditions, however, is close to the threshold of significance and so it should be interpreted with caution. There was not a statistically significant difference noticed between any of the three physical conditions for how social or caring the robot was perceived to be. No single physical condition of the robot made participants significantly happier after hugging it. The addition of the foam and heat proved to be crucial components to significantly increase users' comfort during the hug ($F(2, 58) = 3.17$, $p = 0.049$), with the difference noticed between the Hard-Cold and Soft-Warm conditions.

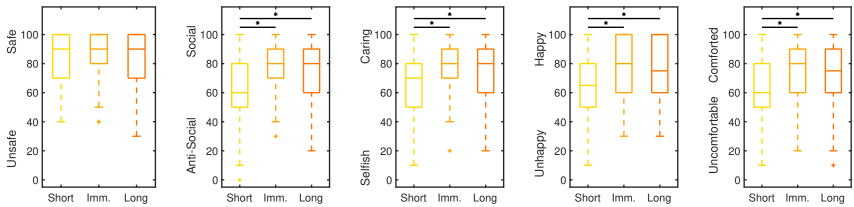


FIGURE 4.10: A comparison of the responses to the survey questions after the last nine hugs, changing behavioral conditions. The top row represents the three levels of the tightness factor (loose, fit, and tight), while the bottom row represents the three levels of the duration factor (short, immediate release, and long). The top and bottom of the box represent the 25th and 75th percentile responses, respectively, while the line in the center of the box represents the median. The lines extending past the boxes show the farthest data point not considered outliers. The + marks indicate outliers.The black lines with stars at the top of the graph indicate statistically significant differences.

4.5.3 *Survey Results from Behavioral Trials*

The last nine hugs varied the behavior of the robot, changing the pressure applied and the duration of the hug. The ratings after each trial were analyzed using a two-way repeated measures analysis of variance and Tukey posthoc tests, as they were for the first three trials. Our data satisfies all the assumptions of a two-way repeated measures ANOVA. The results from the responses to the five questions after each of these hugs can be seen in Fig. 4.10, grouped by level for both tightness and duration factors. No significant interaction effect was found between pressure and duration for any of the five questions asked. The perceived safety of the robot behavior did not significantly change across the nine configurations.

Hug duration played a large role in how anti-social or social users perceived the robot to be, with a significant difference noticed between the too-long and too-short hug, as well as between the just-right and too-short hug ($F(2,58) = 16.057$, $p < 0.005$). No significant difference was found between the too-long and just-right hugs.

The length of the hug also significantly affected how caring people thought the robot was ($F(2,58) = 19.492$, $p < 0.005$). When the robot hugged for too short a time, it was considered more selfish than both the just-right and too-long hug duration, which were thought to be more caring. A significant difference was not noticed between the latter two.

Users were least happy when hugging the robot during the too-short hug, and they were significantly happier when it released on cue or held on for five seconds after they pressed the tactile sensor ($F(2,58) = 8.554$, $p < 0.005$). There was not a significant difference in the participant's happiness between when the robot released on demand or held on too long. The statistical significance between the too long and too short conditions, however, is close to the threshold of significance and so it should be interpreted with caution.

Finally, the participants felt least comfortable when the robot released them too quickly compared to feeling more comforted when it released when they indicated or when it held on too long ($F(2,58) = 7.701$, $p < 0.005$). No difference was found between the last two conditions. The statistical significance between the too long and too short conditions, however, is close to the threshold of significance and so it should be interpreted with caution. No significant differences were found across hug tightness levels.

4.5.4 *Verbal Comments During Experiment*

Another form of data came from verbal comments from participants during the experiment, which were transcribed from the video by an investigator. Five participants (16.7%) commented positively about the softness and warmth of the robot, in comments like "the stomach padding is really nice" and "OOH it's warm!" Three participants (10.0%) responded negatively about the shorter hugs: a common phrase was "did I miss it?" The rest of the participants did not make unprompted verbal comments about the robot. When asked if they were ready for their next hug during the Hard-Cold trial, four users (13.3%) said something like, "no, because it doesn't look very comfortable," and they asked if they had to hug the robot in this condition. A final common negative comment, made by 26 participants (86.7%), was asking "is it working?" or "it's not letting go" during one or more of the intentionally too long hugs. Another common comment made by seven participants (23.2%) before the last trial was disappointment when they realized the experiment was over; one participant (3.3%) said "aw man, last one?" At other points in the study, three people (10.0%) told the experimenter "I like a good, tight hug," while two participants (6.7%) explicitly mentioned their preference of the just right tightness, immediate release duration hug by saying things like "that was the best one yet," or "excellent."

4.5.5 *Free-Response Questions*

The last source of data came from the final comments at the end of the closing survey. When discussing the aspects of the experiment they enjoyed, seventeen people (56.7%) said the hugs in general, eight people (26.7%) mentioned the foam added for softness, and seven people (23.3%) specifically said the warmth improved the enjoyment of the activity. Additionally, eight people (26.7%) discussed the fact that the activity included a robot made it more enjoyable. Four people (13.3%) mentioned that hugging the robot improved their mood, and three (10.0%) noted how much they appreciated the politeness of the robot asking for a hug.

When it came to the aspects of the activity that participants found challenging, there were three main themes. Five of the 30 participants (16.7%) said nothing about this experiment was challenging to them. Eleven people (36.7%) mentioned that the physical aspects of the robotic platform were challenging. These challenges included the large size of the robot base and/or head, which made it difficult to hug, as well as the height restriction that made this activity more challenging for taller individuals, as the PR2 could not match their height. A final aspect that twenty-one participants (70.0%) found challenging was getting the robot to release in several trials. Although we did not ask this question after each individual trial, we believe that these comments largely referred to the three trials when the robot was programmed to recognize that the person had indicated they were ready for the robot to release and intentionally waited 5 seconds before releasing.

When asked why they would or would not want to do this activity, there were three main types of comments. Nineteen people (63.3%) wrote that they were pleasantly surprised by how nice the robot hugs were, and that they would like to do this activity again, making it the most common comment. Next, ten people (33.3%) mentioned that while they did enjoy the activity, they preferred human hugs because they felt that the robot did not understand why the human wanted a hug (their emotional state), and would not react appropriately, e.g., squeeze tighter as they human squeezed tighter, or rub their back when the robot notices they are upset. The last common comment was that five people (16.7%) mentioned they would like to do this activity because they noticed positive improvements in their mood after receiving hugs from the robot.

The final free-response question (what other activities would you like to do with the robot) had five common responses. Ten people (33.3%) mentioned that they would would want to do more hugging activities.

These activities included snuggling, cuddling, receiving pats on the back, back rubs, and massages. Within this topic, six people (20.0%) mentioned that they wanted to have a conversation with a robot and have it be able to determine they were feeling sad, and offer to give them a hug to cheer them up. Ten people (33.3%) said they would like to play games with the robot, six mentioned (20.0%) they'd like to give the robot high-fives, six said they'd like to dance with the robot, and six others (20.0%) talked about how they'd like the robot to be able to talk with them, react to a conversation, and tell them stories. The last common thread among responses was that three people (10.0%) mentioned they'd like a robot to assist in daily tasks.

4.6 DISCUSSION

Our research question was composed of four sub-statements. Three of our four sub-statements were supported by the results of the statistical analysis. When supplemented with verbal and written user feedback, we find the results support all four sub-statements tested. First, H1 hypothesized that users would prefer hugging a cold, soft robot to a cold, hard robot. The addition of the foam to soften the robot improved the comfort of participants during the experiment. Users also considered the soft robot to be much safer than a hard robot. Because of these survey results, as well as the multiple positive comments about the inclusion of the foam, and the number of people who displayed a lack of interest in hugging the Hard-Cold robot, we conclude that people prefer hugging a soft robot to a hard robot.

Next, H2 hypothesized that users would prefer hugging a warm, soft robot to a cold, soft robot. The results from the experiment confirm this hypothesis. The addition of the heat made the most impact on the perceived safety and how comforting the robot was. For these reasons, as well as the numerous positive comments made about the warmth of the robot, both verbally and in writing, we conclude that people prefer hugging a warm robot to a cold robot.

Third, H3 conjectured participants would prefer a robot that hugs them with a medium amount of pressure, rather than hugging too loosely or too tightly. Multiple users verbally mentioned their preference for "really tight hugs," which may help explain some of the results found. Statistical analysis did not find that hug tightness significantly affected answers to any of the questions we asked. We believe that the three studied hug tightness levels were not different enough to cause any measurable effects in these measures. Our too-tight condition was not set very tight out of concern for the safety of our participants. While we initially set out to test too-loose, just right, and too-tight conditions, we believe we actually tested a very loose condition, a slightly loose condition (where the arms of the robot were lightly touching the participant, but not squeezing), and a just-right condition, which our participants verbally told us they liked best. We can thus conclude that humans like a robot that slightly squeezes them during a hug, but we cannot evaluate levels of pressure higher than were tested in the experiment.

Finally, H4 hypothesized that users would prefer a robot that releases them from a hug immediately when they indicated they were ready for the hug to be over, rather than the robot releasing a hug before or after this time. Statistical analysis proved that hug duration played a significant role in the user's experience. In every question with a significant difference noticed (all except for the first question about safety), the too-short hug was less ideal than either the immediate release or the too-long hug. Negative verbal comments were made to the investigator during the trials regarding both the "too short" and "too long" hugs. The number of written comments where the user mentioned it was difficult to get the robot to release indicate discomfort when the robot did not release exactly when they wanted it to. Together, these results make a compelling argument that our users preferred having control over the duration of the hug.

Limitations

While this study was a good starting point, it has its limitations. A clear weakness of the studied approach was the PR2 platform used for this experiment. Because of the intimate nature of this exchange, such a large, bulky robot was a less than ideal choice. It made the encounter uncomfortable for some users and potentially inhibited their enjoyment. This robot's arms also made up its chest, which rotated as it moved its arms. Creating padding that covered the chest area for the entirety of the experiment was therefore difficult. The padding shifted after each hug and had to be repositioned by the experimenter before the next trial. Using a different robotic platform, or developing a new one that is specifically made for such social-physical interactions, would be ideal. Another weakness of this chapter was that we did not equip the robot to match the pressure the human applied; it hugged with three different levels of constant pressure. Due to the delicate question of the safety of telling a robot to hug a person "too tight," we decided to manually pre-program what would be "just right" and "too tight" for each person. A more realistic, but potentially dangerous, way for the robot to hug might be to estimate the applied user pressure and attempt to match it, squeezing tighter and lighter in sync with the user.

The recruitment procedures may also have unintentionally biased our results toward positive assessments because all participants learned the topic of the experiment from the recruitment materials. Additionally, of the participants we recruited, a majority had a technical background. Once the participants arrived, the practice hugs acclimatized participants to the act of hugging a robot, potentially causing higher overall ratings than would be expected without these practice trials. Our study was then conducted in a clinical manner. It is highly unlikely that someone would hug another person twelve times over the course of fifty minutes without speaking to them. We would thus like to conduct another study with a more "in the wild" design, in which human-robot hugs occur in a more natural way, similarly to how humans hug each other. Next, the default configurations used in all the physical and behavioral trials could have influenced our results. Having the robot hug in the "just right" pressure and for a pre-set duration of three seconds during all physical trials may have affected our conclusions about softness and warmth. Having the robot hug in the Soft-Warm condition for all the behavioral trials may have affected our conclusions about pressure and duration. Finally, some of the significant results found could be the result of the demand effect, a commonly stated disadvantage of running a within-subjects study (Charness, Gneezy, & Kuhn, 2012). Recent research, however, suggests that even when participants of aware of the study's purpose, they do not appear to assist researchers (Mummolo & Peterson, 2017). We, nevertheless, acknowledge the possibility that by experiencing all tested conditions, participants could guess the research question and either intentionally or unintentionally bias their responses.

4.7 CONCLUSIONS

This chapter represents an early but important first step in this line of social-physical human-robot interaction research. While robots are being integrated into more tasks with humans, like factory assembly lines and military teams, humans are typically expected to maintain a safe distance outside of the robot's workspace. This thesis aims to bring humans and robots closer, by completely enclosing a human in a robot's arms in a safe and supportive manner. Before more complicated research can be done, it is important to understand the basics of what makes users most comfortable in this new and potentially scary situation. After the positive initial feedback received from this study, it appears that the HRI community could embrace enabling robots to interact physically with humans in typical social interactions they experience with other humans, namely hugs, high fives, games, and dancing. Another implication of this research is that HRI researchers can work to make robots that provide a more enjoyable tactile experience for their users. Every user preferred hugging the robot when it was covered in the soft, fuzzy outfit.

5 HUGGIEBOT 2.0: DESIGN AND EVALUATION OF HUMAN-SIZED HUGGING ROBOT WITH VISUAL AND HAPTIC PERCEPTION

> *"A hug is like a boomerang – you get it back right away."*
>
> — Bil Keane, American Cartoonist

5.1 INTRODUCTION

Our long-term research goal is to determine the extent to which we can strengthen personal relationships that are separated by a physical distance via hugging robots that provide high-quality social touch.

Making a good hugging robot is difficult because it must understand the user's nonverbal cues, realistically replicate a human hug, and ensure user safety. We believe that such robots need multi-modal perception to satisfy all three of these goals, a target no previous system has reached. Some approaches focus primarily on safety, providing the user with the sensation of being hugged without being able to actively reciprocate the hugging motion (Duvall et al., 2016; Teh et al., 2008; Tsetserukou, 2010). Conversely, other researchers focus on providing the user with an item to hug, but that item cannot hug the user back (DiSalvo et al., 2003; Stiehl et al., 2005; Sumioka et al., 2013). Other robotic solutions safely replicate a hug, but they are teleoperated, meaning they have no perception of their user and require an additional person to control the robot any time a user wants a hug (Campbell & Yamane, 2020; Hedayati et al., 2019; Kaplish & Yamane, 2019; Shiomi, Nakata, et al., 2017a; Yamane, Kim, & Alspach, 2017). Finally, some robots have basic levels of perception but are not fully autonomous or comfortable (Block & Kuchenbecker, 2019; Miyashita & Ishiguro, 2004). Therefore, a good hugging robot is challenging because it will need to safely and comfortably understand and respond to a user's non-verbal cues.

To tackle the aforementioned goal of safely delivering pleasant hugs, we propose the first *six tenets of robotic hugging*: a hugging robot should be soft, warm, and sized similar to an adult human, and it should see and react to an approaching user, adjust automatically to that user's size and position while hugging, and reliably respond when the user releases the

FIGURE 5.1: A user hugging HuggieBot 2.0.

embrace. After presenting these tenets and our accompanying hypotheses in Section 5.3, we use the tenets to inform the creation of HuggieBot 2.0, a novel humanoid robot for close social-physical interaction, as seen in Fig. 5.1 and described in Section 5.4. HuggieBot 2.0 uses computer vision to detect an approaching user and automatically initiate a hug based on their distance to the robot. It also uniquely models hugging after robot grasping, using two slender padded arms, an inflated torso, and haptic sensing to automatically adjust to the user's body and detect user hug initiation and termination. HuggieBot 2.0 is the *first human-sized hugging robot with visual and haptic perception for closed-loop hugging*.

In this chapter we seek to validate the four new tenets by conducting two experiments with HuggieBot 2.0. First, we confirmed user preference for the created platform's physical size, visual appearance, and movements through a comparative online study, as described in Section 5.5. We then conducted an in-person study (Section 5.6) to understand how HuggieBot 2.0 and its three new perceptual capabilities (vision, sizing, and release detection) affect user opinions. Section 5.7 discusses the study results, which show that the six tenets significantly improve user perception of hugging robots. Section 5.8 discusses the limitations of our approach and concludes the chapter.

5.2 RELATED WORK

5.2.1 *Using Vision for Person Detection*

One challenge of accurate and safe robotic hugging is detecting a user's desire for a hug. Many researchers solve this problem by using a remote operator to activate the hug (DiSalvo et al., 2003; Sumioka et al., 2013; Yamane, Kim, & Alspach, 2017; Yamazaki et al., 2016). This form of telehug is not a universal approach because it requires a hugging partner to be available at the exact moment a user would like the comfort of a hug. Having the user press a button is a simpler alternative but differs greatly from human-human hugging. Machine vision has been working on the problem of human detection for decades. Early works focus mostly on finding a representative feature set that distinguishes the humans in the scene from other objects. Different methods for the feature extraction have been proposed, such as using Haar wavelets (Oren et al., 1997), histograms of oriented gradients (HOG) (Dalal, Triggs, & Schmid, 2006), and covariance matrices (Tuzel, Porikli, & Meer, 2007), e.g. Several approaches try to combine multiple cues for person detection; for example, Choi et al. (Choi, Pantofaru, & Savarese, 2011) and Vo et al. (Vo, Jiang, & Zell, 2014) e.g., combine the Viola-Jones face detector (Viola & Jones, 2001) and an upper-body detector based on HOG features (Dalal, Triggs, & Schmid, 2006). One method that could allow robots to provide hugs autonomously is to detect an approaching user via computer vision. Human detection has long been of interest in many research fields, including autonomous driving (Yurtsever et al., 2020), surveillance and security (Paul, Haque, & Chakraborty, 2013), computer vision (Viola & Jones, 2004), and human-robot interaction (Vo, Jiang, & Zell, 2014). In computer vision, deep-learning-based systems relying on convolutional neural networks (CNNs) are often used for person detection. However, the computational cost of these detection pipelines is very high, so they are not suited for use on a real-time human-robot interaction platform. The recent decrease of the computational cost of improved models, e.g., (Liu et al., 2016), has facilitated their adaptation to robot platforms. Hence, we integrate such a model into our pipeline so that HuggieBot 2.0

can recognize an approaching user to initiate a hug with minimal on-board computational load.

5.2.2 *Hugging as a Form of Grasping*

Once the user arrives, safely delivering a hug is challenging for robots because users come in widely varying shapes and sizes and have different preferences for hug duration and tightness. No existing hugging robots are equipped to hug people adaptively. We propose looking to the robotics research community to find a solution. Grasping objects of varying shape, size, and mechanical properties is a common and well-studied problem, e.g., (Barber et al., 1986; Costanzo, De Maria, & Natale, 2020; Ma, Spiers, & Dollar, 2016; Romano et al., 2011). Therefore, we look at hugging as a large-scale two-fingered grasping problem, where the item to be grasped is a human body. For example, the BarrettHand, a commercially available three-fingered gripper, automatically adjusts to securely grasp widely varying objects by closing all finger joints simultaneously and then stopping each joint individually when that joint's torque exceeds a threshold (Townsend, 2000).The robot arms used for HuggieBot 2.0 have torque sensors at every joint, making this torque-thresholded grasping method an ideal way to achieve a secure embrace that neither leaves air gaps nor applies excessive pressure to the user's body. Torque sensors can also enable the robot to feel when the user wishes to leave the embrace.

5.3 HUGGING TENETS AND HYPOTHESES

We propose six tenets to guide the creation of future hugging robots. The first two were validated by the study reported in the previous chapter (Block & Kuchenbecker, 2019), and the other four are proposed and validated in this chapter. We believe that a hugging robot should:

T1. be soft,

T2. be warm,

T3. be sized similar to an adult human,

T4. visually perceive and react to an approaching user, rather than requiring a triggering action such as a button press by the user, an experimenter, or a remote hugging partner, T5. autonomously adapt its embrace to the size and position of the user's body, rather than hug in a constant manner, and

T6. reliably detect and react to a user's desire to be released from a hug regardless of their arm positions.

Building on the previously described research, this chapter seeks to evaluate the extent to which the six tenets benefit user perception of hugging robots. Specifically, we aim to test the following three hypotheses:

H1. When viewing from a distance, potential users will prefer the embodiment and movement of a hugging robot that incorporates our four new tenets over a state-of-the-art hugging robot that violates the tenets.

H2. Obeying the four new tenets during physical user interactions will significantly increase the perceived safety, naturalness, enjoyability, intelligence, and friendliness of a hugging robot.

H3. Repeated hugs with a robot that follows all six tenets will improve user opinions about robots in general.

5.4 SYSTEM DESIGN AND ENGINEERING

We introduce a new human-sized hugging robot with visual and haptic perception. This platform is designed to have a friendlier and more comfortable appearance than previous state-of-the-art hugging robots. Building on feedback from users of prior robots (Section 5.2) and HuggieBot 1.0 (Chapter 4), we focused on six areas for the design of this new, self-contained robot: the frame, arms, inflated sensing torso, head and face, visual person detection, and software architecture.

5.4.1 *Frame*

HuggieBot 2.0's core consists of a custom stainless steel frame with a v-shaped base. The robot's height can be manually adjusted if needed, and the shape of the base allows users to come very close to the robot, as seen in Fig. 5.1. The user does not need to lean over a large base to receive a hug, unlike the PR2-based hugging robot (Block & Kuchenbecker, 2019). The v-shaped base also increases the safety and stability of the robot by acting to counteract any leaning force imparted by a user. This large base and counterweight ensure that even a large user approaching at a high speed intending to make an incorrect impact with the robot will not flip it over, inflict injury upon themselves, or cause damage to the robot.

5.4.2 *Arms*

Two 6-DOF Kinova JACO arms are horizontally mounted to a custom stainless steel bracket attached to the top of the metal frame. To create a more approachable appearance, the grippers of the JACO arms were removed, and large padded mittens were placed over the wrist joints that terminate each arm. The arms are controlled by commanding target joint angles; movement is quiet, and the joints are not easily backdrivable when powered. The torque sensors at each joint are continually monitored so that hugs can be automatically adjusted to each user's size and position. The second joint (shoulder pan) and third joint (elbow flex) on each arm stop closing individually when they surpass a torque threshold, which we empirically set to 10 Nm and 5 Nm, respectively. The joint torques are also used to detect when a user is pushing back against the arms with a torque higher than 20 Nm, indicating their desire to be released from a hug. To create a comfortable and enjoyable tactile experience for the user, we covered the arms in soft foam and a sweatshirt.

5.4.3 *HuggieChest: Inflatable Torso*

We developed a simple and inherently soft inflatable haptic sensor to serve as the torso of our hugging robot, as pictured in Fig. 5.2. The torso was constructed by both heat sealing and gluing (with HH-66 vinyl cement) two sheets of PVC vinyl to create an airtight seal. This chest has one chamber located in the front and another in the back. There is no airflow between the two chambers. Each chamber has a valve from an inflatable swim armband to inflate, seal, and deflate the chamber. Inside the chamber located on the back of the robot is an Adafruit BME680 barometric pressure sensor and an Adafruit electret microphone amplifier MAX4466 with adjustable gain. Both sensors were secured in the center of the chest on the inner wall of the chamber. The two sensors are connected to a single Arduino Uno micro-controller outside the chamber. The microphone and pressure sensor are sampled at 45 Hz, and the readings are sent over serial to the HuggieBot 2.0 computer for real-time processing. We originally tested with the same sensing capabilities in both chambers but did not find the information from

FIGURE 5.2: The inflated HuggieChest when lying flat. The two air chambers form the front and back of the robot's torso. One barometric pressure sensor and one microphone are located inside the chamber that belongs on the robot's back (shown at left in this photo).

the front chamber to be very useful; thus, no data are collected from the front chamber. This novel inflatable haptic sensor is called the HuggieChest.

The HuggieChest's shape was created by following a pattern of a padded vest that goes over the wearer's head and is secured with a belt around the waist. Because the HuggieChest is heat-sealed at the shoulders to stop airflow between the chambers and allow the chest to bend once inflated, the robot is unable to feel contacts in these locations. The HuggieChest is placed directly on top of the metal frame of HuggieBot 2.0. On top of the inflatable torso, we put two Thermophore MaxHeat Deep-Heat Therapy heating pads (35.6 cm $\times$ 68.6 cm), which are attached together at one short edge with two shoulder straps. From Trovato et al. (2016) we learned that softness alone is not enough for a hugging robot; people are more receptive to hugging a robot that is wearing clothing, so we placed suitable clothes on HuggieBot 2.0. A sweatshirt is placed on top of the heating pads to create the final robot torso.

5.4.4 *Head and Face*

We designed and 3D-printed a head to house a Dell OptiPlex 7050 minicomputer that controls the robot, a face screen, an RGB-D camera, the Arduino from the HuggieChest, a wireless JBL speaker, and cables. As shown in Fig. 5.3, the head splits into two halves with a rectangular plate on each side that can be removed to access the inside. The final piece of the head is the front-facing frame, which secures the face screen and camera. The face

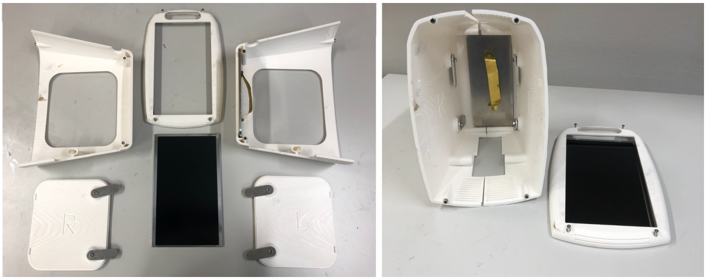


FIGURE 5.3: The head disassembled and partially assembled.

screen is an LG LP101WH1 Display 10.1″ LCD screen with a 1366 × 768 resolution in portrait orientation. The screen displays faces based on designs created and validated for the Baxter robot (Fitter & Kuchenbecker, 2016a).The size and orientation of the faces were altered to fit HuggieBot 2.0's screen.

5.4.5 *Vision and Person Detection*

We use a commercially available Intel Realsense RGB-D camera with custom software to recognize an approaching person and initiate a hug. To this end, we integrate a deep-learning-based person detection module into our pipeline. The module consists of two parts. First, our software recognizes an approaching person using an open-source Robot Operating System (ROS) integration (Odabasi, 2017) of Tensorflow's object detection library, which is based on the SSD mobilenet model (Liu et al., 2016). In the next step, we utilize the camera's depth sensor to estimate the distance of the person to the robot. We use a sliding window to ensure a person is actively approaching the robot; we observe the distance measured from the depth sensor, which can be noisy, and check whether the mean distance decreases. This strategy prevents undesired hugs in case a person walks away from the robot. Once the person is actively approaching the robot, we initiate a hug as soon as a tuned distance threshold of 2.45 m is passed. This experimental threshold was selected as it informs the robot the person is attempting to move from the social space into the robot's personal space (Hall et al., 1968). Researchers have also looked in-depth at the use of social space in HRI. One group found that users underestimate the distance to the location of robot peak performance and that people adjust their proxemic preferences to be near the perceived location of robot peak performance (Mead & Matarić, 2015). This group further analyzed individual, physical, and psychophysical factors in real-time that contribute to proxemic preferences (Mead & Matarić, 2017). They trained Hidden Markov Models on both physical and psychophysical features to identify the initiation and termination of a social interaction by a user, which outperformed those only trained on physical features (Mead, 2016; Mead & Matarić, 2017). While all the studies conducted in this thesis took place in a single room in a controlled environment, future studies could potentially take place in a less controlled environment, with more noise in the background. Implementing a more sophisticated model of determining social space like this would be a better approach to determine an approaching user wanting a hug to avoid hand tuning to the environment. Finally, we have also implemented the ability for the robot's camera to scan QR codes from an app and hug in a customized manner, but this feature is not evaluated in this thesis.

5.4.6 *Robot Software Architecture*

The robot is controlled via ROS Kinetic. Each robot arm joint has both angle and torque sensors. A PID controller is used to control each joint angle over time. The robot arms begin by moving to a home position. The camera module starts and waits for an approaching user. Upon detection, the robot asks the user, "Can I have a hug, please?" as in Block and Kuchenbecker (2018), while the robot's face changes to an opening and closing mouth. The specific hug it is supposed to run (with or without haptic sizing and release) is executed by commanding each joint to move at a fixed angular velocity toward a predetermined goal pose. For hugs without haptic sizing, the robot hugs in a one-size-fits-most manner, where the robot's second and third joints each close by 20°. This pose was large enough such that it did not apply high forces to the bodies of any of our users; it was not adjusted for different users. For hugs with haptic sizing, the robot arms move toward a pose sized for a small user; we continually monitor each joint torque and stop a joint's movement if it exceeds the pre-set torque threshold. This method leads to automatic adjustment to the user's size (T5).

The Arduino communicates the microphone and pressure sensor data from inside the back chamber of the HuggieChest to ROS over serial at 45 Hz. This data stream is analyzed in real time. The program first determines the ambient pressure and noise in the chamber by averaging the first 20 samples to create a baseline that accounts for different levels of inflation and noise. The chamber detects the user beginning to hug when the chamber's pressure increases by 50 kPa above the baseline pressure. Contact is determined to be broken, thus indicating that the user wants to be released, when the pressure returns to 10 kPa above the baseline pressure. The measured torques from the robot's shoulder pan and elbow flex joints are monitored continually during a hug. A haptic release is also triggered when any of these torques surpasses a threshold limit of 20 Nm. For a timed hug, rather than detecting the instant at which the user wants to be released, the robot waits 1 second after the arms fully close before releasing the user, so it is apparent to the user they are not in control of the duration of the hug. Overall, our proposed method of closed-loop hugging works on a higher level of abstraction than the low-level control, i.e., including both visual and haptic perception in the loop of the hugging process. The robot's haptic perception is two-fold: adjusting to the size of the user and sensing when he/she wants to be released.

5.5 ONLINE USER STUDY

We ran an online study to get feedback from a broad audience on the embodiment and movement of our robot as part of our user-centered design process, and to compare it to the PR2-based HuggieBot 1.0 (Block & Kuchenbecker, 2019). The main stimuli were two videos from Block and Kuchenbecker (2019) along with two newly recorded videos of people hugging HuggieBot 2.0 with matched gender, enthusiasm, and timing; these videos are included as supplementary material for Block, Christen, et al. (2021). This study was conducted under the Haptic Intelligence Department's framework agreement with the Max Planck Society Ethics Council as protocol F002B.

5.5.1 *Participants*

All participants for the online study were non-compensated English-speaking volunteers recruited via emails and social media. A total of 117 participants took part in the online survey: 42.7% male, 56.4% female, and 0.9% who identify as other. The participants ranged in age from 20 to 86 (M = 37.5, SD = 16.75). The majority of respondents indicated they had little (30.7%) or no experience (43.6%) interacting physically with robots.When asked in more detail about their level of experience with robots most indicated they were novice (36%) or beginners (36%). The last question was only asked to the last 97 participants, rather than all 117, and will be explained below.

Question
This hug made the robot seem (unfriendly – friendly)
This robot behavior seemed (unsafe – safe)
This hug made the robot seem socially (stupid – intelligent)
This hug interaction felt (awkward – enjoyable)
This robot behavior seemed (unnatural – natural)

Table 5.1: The questions participants answered after viewing or experiencing robot hugs.

5.5.2 *Procedures*

After someone reached out to the experimenter and indicated interest in participating in the online study, the experimenter sent the person an informed consent document by email. The participant read it thoroughly, asked any questions, and signed it and sent it back to the experimenter only if they wanted to participate. At this point, the user was assigned a participant number and sent a unique link to the online survey.The experimenter alternated which survey was given to the users, which varied the presentation order of the two robots shown.

First, participants filled out their demographic information, including their gender, age, level of extroversion, and country of origin. Included in this part of the survey, we asked about the number of hugs they typically exchanged before COVID-19, how much they typically enjoy hugs, their level of experience interacting with robots, and their experience physically interacting with robots. Next, they were shown two videos of an adult (one male and one female) hugging a robot labeled "Robot A". Robot A was HuggieBot 2.0 for half of the participants and HuggieBot 1.0 (Block & Kuchenbecker, 2019) for the other half. One video showed a male hugging the robot, while the other showed a female. The order of these videos did not change. Users could watch these videos as many times as they liked before answering several questions. Users described their first impressions of the robot they saw. Then, they answered the questions shown in Table 5.1 on a 5-point Likert scale. Afterwards, there was an optional space for additional comments. Next, they were shown two videos of people hugging the other robot labeled "Robot B" under the same conditions as the first videos. The experimenters worked to match the timing, speed, and voice

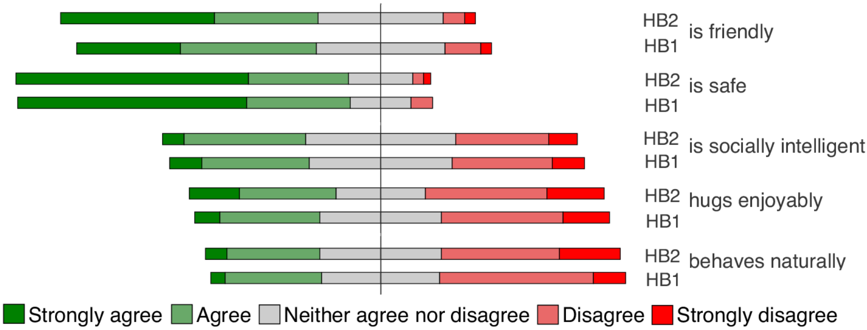


FIGURE 5.4: The responses to the five questions asked after users watched two videos of people hugging HuggieBot 2.0 (HB2) and HuggieBot 1.0 (HB1).

of the robot, as well as the level of excitement of the users in the video so as not to sway participants by these factors. Users answered the same questions for Robot B as they did for Robot A. Finally, since the videos were shot from behind the robot, users were shown frontal images of both robots posed in a similar manner. Participants were then asked "In what ways is Robot A better than Robot B?" and "In what ways is Robot B better than Robot A?" Then, they were asked to select which robot they would prefer to hug, Robot A or B, and why.

After the first 40 participants, we noticed that several users were commenting on the purple fuzzy appearance of the PR2, rather than on the robots themselves. Therefore, we added a new section at the end of the survey, which was completed by the remaining 77 participants. This new section showed photographs of Robot A and Robot B plus two additional photographs showing HuggieBot 2.0 in different clothing. In one of the new photographs, HuggieBot 2.0 wore a fuzzy purple robe similar to the fabric cover on the PR2, and in the other it wore its gray sweatshirt over this same fuzzy purple robe. Participants were asked which of these four robots they would prefer to hug and why.

Results

The responses to the five Likert-style questions from Table 5.1 can be seen in Fig. 5.4. For all statistical analyses, we applied a Bonferroni alpha correction to $\alpha = 0.05$ to determine significance and to account for the multiple comparisons. Because the data from these questions were non-parametric,

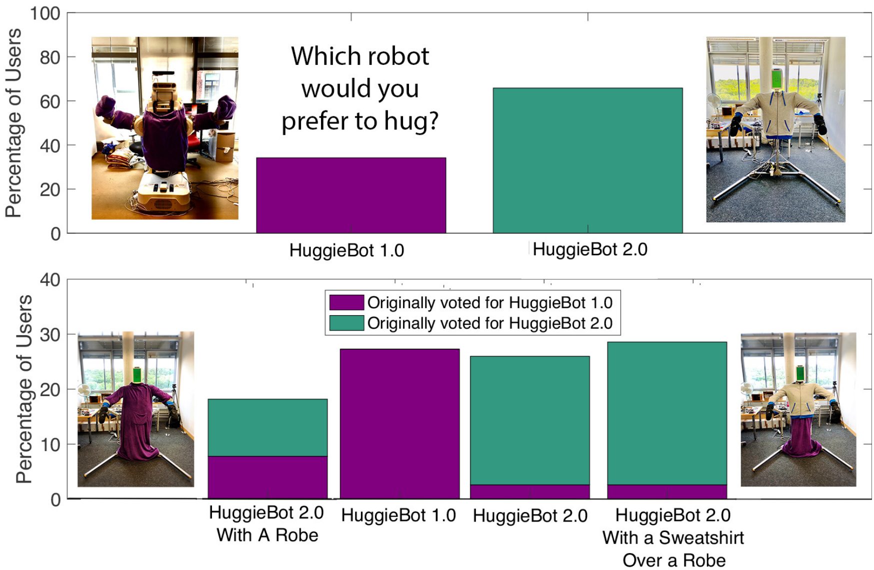


FIGURE 5.5: The breakdown of preferences when users had two choices (top) and four choices (bottom), with the associated images. The colors of the second plot show which robot the user preferred in the first selection round.

we conducted a Wilcoxon signed-rank test. No statistically significant differences were found between the responses to any of these questions for the two robots. The responses to the robot friendliness question approached significance ($p = 0.061$), with users tending to indicate HuggieBot 2.0 seemed slightly friendlier.

The responses to the first and second rounds of voting for which robot users would prefer to hug can be found in Fig. 5.5. We ran a Wilcoxon signed-rank test for the first round of voting and determined users would significantly prefer to hug our new robot over HuggieBot 1.0 ($p < 0.001$). In the second voting round, no significant preference was found between any of the four options, indicating HuggieBot 2.0 was preferred over HuggieBot 1.0 approximately 3:1. We ran several one-way analyses of variance (ANOVA) to see if participant gender, robot presentation order, or participant level of extroversion had a significant effect on which robot the user selected. No significance was found in any of these cases.

5.5.3 *Changes to Platform*

After analyzing the results of the online study and reviewing the user feedback, we found several areas to improve our hugging robot. Some users found the initial HuggieBot 2.0 voice off-putting, so we changed the robot's voice to sound less robotic. We made the purple robe and the sweatshirt the robot's permanent outfit, as it had the highest number of votes and received many positive comments. Even users who did not select our robot with the robe outfit indicated in their comments that they liked the idea of covering up the metallic components. We changed the color of the robot's face from its initial green to purple to match the robe and create a more polished look. Several users commented on the slow speed of HuggieBot 2.0's arms. Since the arm joints cannot move faster than the maximum angular velocity specified by the manufacturer, we instead moved their starting position closer to the goal position to reduce the time they need to close.

5.6 IN-PERSON USER STUDY

The goal of the in-person study was to evaluate our updated robotic platform and the four new hugging tenets that drove its design. This study was also approved under the Haptic Intelligence Department's framework agreement with the Max Planck Society Ethics Council under protocol F006B.

5.6.1 *Participants*

The recruitment methods for the in-person study were the same as for the online study. Participants not employed by the Max Planck Society were compensated 12 euros. A total of 32 users participated in the in-person study: 37.5% male and 62.5% female. Our participants ranged in age from 21 to 60 (M = 30, SD = 7) and came from 13 different countries.

5.6.2 *COVID-19 Health Safety Measures*

Given the COVID-19 global pandemic, additional health and safety precautions were taken during the running of our in-person study. All participants were required prior to their arrival on our campus to fill out a form verifying they had not been in a risk region in the last 14 days, that they had no current symptoms, nor had they been in contact with anyone who had symptoms. If a participant had been in a risk zone within the last 14 days, but had received two negative COVID-19 tests and submitted them to us, they were allowed on campus. All participants were then required to thoroughly wash their hands or use hand sanitizer supplied in the experiment room. While the participant did not wear a mask during the experiment so their facial expressions could be captured on film and later analyzed, the experimenter wore a mask when coming closer than 1.5 meters to the participant. The windows of the room were always kept open for proper ventilation and air circulation. The robot's sweatshirt was replaced between each user to minimize the risk of potential infection. Additionally, between each user, the robot's head and the tablet users answered the survey questions on were wiped down with anti-viral sanitizer.

5.6.3 *Inclusion/Exclusion Criteria*

All users were required to be over 18 years old. All users were also required to speak and understand English well (roughly at B2 level, though it was not checked). Users who were pregnant or other vulnerable people were excluded from participation. Additionally, no one with uncorrected motor or vision disabilities was allowed to participate. Finally, no one with COVID-19 symptoms or recent exposure to someone with the symptoms or recent visit to a risk zone was able to participate in the study.

5.6.4 *Procedures*

After confirming their eligibility given the exclusion criteria, users scheduled an appointment for a 1.5-hour-long session with HuggieBot 2.0. Upon arrival, the experimenter explained the study, and the potential participants read the informed consent document and asked questions. If he/she still wanted to participate, the user signed the consent form and the video release form, at which point the video cameras were turned on to record the experiment.

Users began by filling out a demographics survey on a computer. Next, the investigator introduced the robot as the personality "HuggieBot" and explained its key features, including the emergency stop. The experimenter explained how the trials would work and how the participant should be prepared to move. She also explained the two different ways to initiate a hug (walking, key press) and the three different ways to be released from a hug (release hands, lean back, wait until robot releases). At this point the user filled out an opening survey to document their initial impressions of

Hug Conditions
No Vision, No Sizing, No Release Detection
No Vision, No Sizing, Release Detection
No Vision, Sizing, No Release Detection
No Vision, Sizing, Release Detection
Vision, No Sizing, No Release Detection
Vision, No Sizing, Release Detection
Vision, Sizing, No Release Detection
Vision, Sizing, Release Detection

TABLE 5.2: The eight hugging conditions users experienced in the in-person study.

the robot; participants rated on a sliding scale from 0 (disagree) to 10 (agree) how much they agreed with the statements found in Table 5.4. Note that these questions were asked before the user had any experience physically hugging the robot. Next, the user performed practice hugs with the robot to acclimate to the hug initiation methods and the timing of the robot's arms closing. A participant was allowed to perform as many practice hugs as desired, verbally indicating to the experimenter when they were ready to begin the experiment. On average users did 2 or 3 practice hugs, but users taller than the robot (1.75 m) averaged 5 or 6 practice hugs because it took more time for them to find the most comfortable arm positions. The eight hugging conditions that made up this experiment are all three possible pairwise combinations of our three binary factors (vision, sizing, and release detection), as shown in Table 5.2. The video associated with Block, Christen, et al. (2021) shows a hugging trial from each of the eight conditions.

We used an 8×8 Latin square to counter-balance any effects of presentation order (Grant, 1948). The formula to create an $n \times n$ balanced Latin square can be seen in Table 5.3. For this experiment, we used an 8×8 Latin square. In order to ensure a fully counter-balanced presentation order with complete Latin squares, we recruited 32 participants. After each hug, the participant returned to the computer and answered six questions. The first question was a free-response asking for the user's "first impressions of this interaction." Then, the participant used a sliding scale from 0 to 10 to answer the five questions found in Table 5.1, which were the same questions

1	2	n	3	$n-1$	4	...
2	3	1	4	n	5	...
3	4	2	5	1	6	...
...	...	...	...	...	....	...
$n-1$	n	$n-2$	1	$n-3$	2	...
n	1	$n-1$	2	$n-2$	3	...

Table 5.3: Latin square example to counter-balance presentation order of hugs.

as in the online study. A participant could request to experience the same hug again if needed. After experiencing all eight hug conditions, the participants experienced an average of 16 more hugs, during which they contacted the robot's back in different ways and received various robot responses. Data were collected for these additional hugs, but they will not be analyzed in this chapter due to space constraints. At the end of the experiment, the user answered the same questions from the beginning of the study (Table 5.4). Finally, users could provide additional comments at the end.

All slider-type questions in the survey were based on previous surveys in HRI research and typical Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaires (Heerink et al., 2009; Weiss et al., 2008). The free-response questions were designed to give the investigators any other information the participant wanted to share about the experience. A within-subjects study was selected for this experiment because we were most interested in the differences between the conditions, rather than the overall response levels to a robot hug. We also preferred this design for its higher statistical power given the same number of participants compared to a between-subjects study (de Winter & Dodou, 2017).

5.6.5 *Results*

This in-person study was the first robustness test of the fully integrated HuggieBot 2.0 system. Each user experienced a minimum of 24 hugs during the study, plus practice hugs. With 32 total participants, the robot executed more than 850 successful hugs over the course of the entire study, sometimes giving 200 hugs in one day.

Questions
I feel understood by the robot
I trust the robot
Robots would be nice to hug
I like the presence of the robot
I think using the robot is a good idea
I am afraid to break something while using the robot
People would be impressed if I had such a robot
I could cooperate with the robot
I think the robot is easy to use
I could do activities with this robot
I feel threatened by the robot
This robot would be useful for me
This robot could help me
This robot could support me
I consider this robot to be a social agent

TABLE 5.4: The fifteen questions asked in the opening and closing questionnaires.

For all statistical analyses, we applied a Bonferroni alpha correction (to account for 15 multiple comparisons) to $\alpha = 0.05$ to determine significance. We use Pearson's linear correlation coefficient, ρ, to report effect size. Box plots of the responses to the opening and closing survey questions from Table 5.4 are shown in Fig. 5.6. In this study, answers were submitted on a continuous sliding scale, so a paired t-test comparison of the opening and closing survey was conducted. We found that users felt understood by ($p = 0.0025$, $\rho = 0.57$) and trusted the robot more ($p < 0.001$, $\rho = 0.70$) after participating in the experiment. Users also felt that robots were nicer to hug ($p < 0.001$, $\rho = 0.76$)

The responses to the five questions asked after each hug can be seen grouped by the presence and absence of each of the three tested factors (vision, sizing, and release) in Fig. 5.7. These responses were analyzed using three-way repeated measures analysis of variance via the built-in MATLAB function `ranova`; our data satisfy all assumptions of this analytical

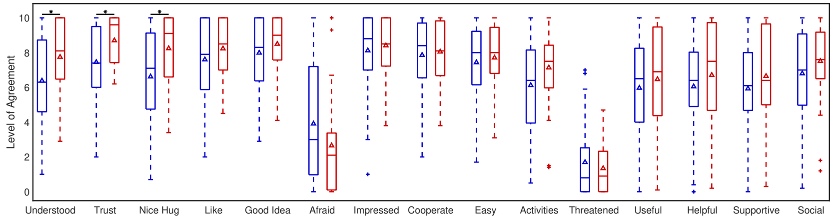


FIGURE 5.6: A comparison of the responses to the opening (blue) and closing (red) surveys. The top and bottom of the box represent the 25th and 75th percentile responses, respectively, while the line in the center represents the median, and the triangle indicates the mean. The lines extending past the boxes show the farthest data points not considered outliers. The + marks indicate outliers. The black lines with stars at the top of the graph indicate statistically significant differences.

approach. No significant improvements were noticed in the perceived safety of the robot in any of the tested conditions, as the robot was consistently rated highly safe. The automatic size adjustment significantly increased users' impressions of the naturalness of the robot's movement ($F(1,31) = 25.192$, $p < 0.001$, $\rho = 0.4158$). Users found the robot's hug significantly more enjoyable when it adjusted to their size ($F(1,31) = 70.553$, $p < 0.001$, $\rho = 0.3610$). Automatic size adjustment to the user caused a significant increase ($F(1,31) = 25.102$, $p < 0.001$, $\rho = 0.4258$) in the perceived intelligence of the robot. Finally, the robot was considered significantly friendlier when it adjusted to the size of the user ($F(1,31) = 84.925$, $p < 0.001$, $\rho = 0.3205$). In summary, haptic hug sizing significantly affected every aspect except safety. Trials that included visual perception trended slightly positive but were not significantly different for any five of the investigated questions; small positive trends for haptic release were also not significant.

Eight users (25%) verbally stated and wrote about their preference for not having to push a button to activate a hug. Out of the 256 distinct hug response surveys (32 users, 8 surveys per user), the physical warmth of the robot was positively mentioned 100 times (39%), further validating that physical warmth is critical to pleasant robot hugs (T2) (Block & Kuchenbecker, 2019). These positive comments were most commonly seen in the conditions with automatic hug sizing, presumably because the increased contact with the robot torso made the heat more apparent. Additionally, we observed that our participants used a mixture of pressure release and torque release to indicate their desire to end the hugs in the study. 17 users (53%) voiced their preference for the haptic release hugs, saying when the robot

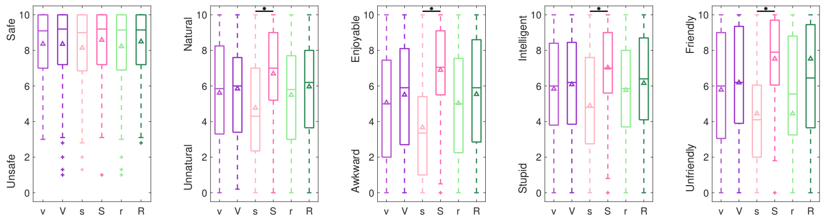


FIGURE 5.7: A comparison of the responses to the survey questions after each hug, grouped by factors. The purple colors represent without vision (v) and with vision (V), the pink colors represent without sizing (s) and with sizing (S), and the green colors represent timed release (r) and haptic release (R). The lighter shade of each color indicates the level without the factor, while the darker color indicates when the factor is present.

released before they were ready (in hugs with a timed release), "he didn't want to hug me!" or that the hug was "too short!" Interestingly, the hug condition where all three perceptual factors were present had the highest number of positive comments. 31 out of 32 users (96.8%) commented that this condition was the most "pleasant interaction," "natural," "friendly," or "fun."

5.7 DISCUSSION

Our three hypotheses were largely supported by the results. First, H1 hypothesized that when viewing from a distance, potential users will prefer the embodiment and movement of a hugging robot that incorporates our four new tenets over a state-of-the-art hugging robot that violates these tenets. The online study found that users significantly preferred HuggieBot 2.0 over HuggieBot 1.0; we believe our new robot was preferred because it obeys all six tenets. Written comments from the online community mention the "large," "hulking," and "over-powering" PR2 robot as unnerving when compared to the size of the user. In comparison, our robot is considered "nice" and "friendlier." Several users also wrote comments on how the people in the PR2 videos had to push a button on the robot's back, which seemed "unnatural", whereas the HuggieBot 2.0 release seemed more "intuitive". We did our best to match the videos of our new robot to the pre-existing videos of the PR2 so as not to bias the online viewers. These videos included but could not showcase visual hug initiation and haptic size adjustment. Based on the strong preference for our hugging robot in our carefully controlled online study, we conclude that users do prefer the embodiment and movement of a hugging robot that obeys our four new tenets over a state-of-the-art hugging robot that violates most of them.

H2 conjectured that obeying the four new tenets during physical interactions with users will significantly increase the perceived safety, naturalness, enjoyability, intelligence, and friendliness of a hugging robot. We found that the haptic perception tenets had the greatest effects on these aspects of the robot, with haptic sizing positively affecting many responses and both haptic sizing and haptic release garnering positive comments. The lack of significant effects of visual initiation does not match the comments that users prefer the interaction when the robot recognizes their approach, rather than them having to push a button to initiate the hug. It is possible that users might not have included the button pushing in their rating of the hug as we simply asked users to "rate their experience with this hug" and did not explicitly tell them to include the hug initiation. Users might also have been confused that they had to walk towards the robot to initiate a hug, and then the robot would ask "Can I have a hug, please?" We chose to have the robot say the same phrase for both initiation methods to minimize variables, but as the user was initiating the hug, it may have made more sense for the robot to say something else or not speak.

We also believe there is room for improvement of the visual perception of our robot, which could contribute to higher ratings of the five questions. Currently, our perception of the user is based solely on their approach. To take perception even further, we believe hugs would be more comfortable if the robot could adjust its arm poses to match the approaching user's height and arm positions. Our taller users found the robot hugged them too low, and our shorter users found the opposite. Adjusting to user height would more fully obey T4 (visual perception) and therefore should be more acceptable to users. While the torso of the robot and dual release methods ensure our robot follows T6 of reliably releasing the user, a robot that could adjust its arm positions to the reciprocating pose of the user could greatly strengthen the user's impression of the robot's visual perception and improve the user opinion of the robot. We concede that our robot's rudimentary visual perception contributed to our lack of finding significant differences in the areas we investigated when testing with and without this factor.

Finally, H3 hypothesized that repeated hugs with a robot that follows all six tenets will improve user opinions about robots in general. We asked the same opening and closing survey questions as Block and Kuchenbecker

(2019), with similar results. Both studies' users felt more understood by, trusted, and thought that robots were nicer to hug after participating. HuggieBot 1.0's users also liked the presence of the robot more afterwards, found the robot easier to use, and viewed it as more of a social agent after the experiment, although these findings were reported without any statistical correction for multiple comparisons. The PR2 robot used in that experiment is significantly larger than an adult human, which violates T3. This domineering physical presence, therefore, contributed to lower initial ratings for users liking the presence of the robot, their perceived ease of use of the robot, and viewing the robot as a social agent. Our new robot, whose physical stature obeys the first three tenets, received higher initial ratings in these categories. HuggieBot 2.0 appeared as a friendly social agent from the beginning, and prolonged interaction with it confirmed these high initial impressions, which is why we did not find any significant differences for these questions in our study. As first impressions are often critical to determine whether a user will interact with a robot, here we see that it is important to obey the tenet prescribing the physical size of a robot. Therefore, we conclude that a robot that follows the six tenets does indeed improve user opinions about robots in general. A positive impression of the robot is crucial because it will make users more willing to receive a robot hug, and thus more likely to receive these health benefits when they can't receive them from other people.

5.8 LIMITATIONS AND CONCLUSION

This study represents an important step in understanding intimate social-physical human-robot interactions, but it certainly has limitations. Due to COVID-19, the first study relied solely on videos and images, rather than the participants physically interacting with robots. Since we do not have access to a PR2, this online study enabled a fair comparison between our new platform and a previously well-rated hugging robot. Watching other people hug a robot is how users will decide if they also want to interact with a hugging robot in the wild.

One weakness of our new platform, HuggieBot 2.0, was the slow speed of the Kinova JACO arms. We selected these arms because of their inherent safety features; however, the distance the arms had to travel made their slow speed obvious and caused a long delay after hug initiation. When users started the hug with a button press, they could wait before walking to the robot for better timing. With visual hug initiation, the users were required to walk before the arms began moving, which resulted in many awkwardly waiting in front of the robot for the arms to close. Related to this limitation is the size of the room in which we conducted the experiment. A larger room would have let us set the threshold distance farther back to accommodate the speed of the arms. Users would thus not have had to wait at the robot before hugging or hug for a prolonged time. Both of these limitations could have contributed to the lack of significant differences between the hugs with and without visual perception.

Another limitation is the self-selection bias of our participants. For transparency, we advertised the experiment as a hugging robot study. While we succeeded in recruiting a diverse and largely non-technical audience to make our results as applicable to the general public as possible, we nevertheless acknowledge that users who chose to participate in the study were interested in robots. Because we did not hide the nature of our study, we did not have any participants who refused to hug the robot, as might occur in a more natural in-the-wild study design.

This chapter took a critical look at state-of-the-art hugging robots, improved upon their flaws, built upon their successes, and created a new hugging robot, HuggieBot 2.0. We also propose to the HRI community six tenets of hugging that future designers should consider to improve user acceptance of hugging robots. During times of social distancing, the consequences of lack of physical contact with others can be more damaging and prevalent than ever. If we cannot seek comfort from other people due to physical distance or health or safety concerns, it is important that we seek other opportunities to reap the benefits of this helpful interaction.

6

HUGGIEBOT 3.0: DESIGNING AUTONOMOUS HUGGING ROBOTS WITH INTRA-HUG GESTURES

A hug is always the right size!

— Winnie the Pooh

6.1 INTRODUCTION

Hugs that last more than three seconds often include intra-hug gestures, like squeezes and rubs (Fig. 6.1). These types of embraces create a close physical exchange between the two participants and confer additional benefits from the increased deep pressure touch (Nagy, 2011).

Accurately replicating a human hug is a difficult problem because it requires *real-time adaptation to a wide variety of users*, *close physical contact*, and *quick, natural responses to intra-hug gestures performed by the user*. In the past, researchers have avoided tackling these challenges by providing a huggable device that does not actively hug the user back, thereby entirely avoiding the challenges of reciprocating a hug (DiSalvo et al., 2003; Stiehl et al., 2005; Sumioka et al., 2013; Yamazaki et al., 2016). Others have chosen to create robots that hug in a "one-size-fits-most" model (Genz, 2007; Hedayati et al., 2019; Shiomi, Nakata, et al., 2017a, 2017b; Tsetserukou, 2010). Another set of researchers adjusted the robot to each specific user prior to experimentation, thereby avoiding the challenge of real-time adaptation (Block & Kuchenbecker, 2018, 2019). The previous chapter introduced HuggieBot 2.0 as the first robot that uses visual and haptic perception to deliver closed-loop hugging that adapts to the circumference of the user and their preferred hug timing (Block, Christen, et al., 2021); however, this robot could not perceive or respond to intra-hug gestures, and user testing revealed other limitations.

To create a robot that can autonomously deliver pleasant, natural-feeling hugs, we conducted a Wizard-of-Oz user study (action-response elicitation study) with HuggieBot 2.0 to collect data on intra-hug gestures; Section 6.2 explains the methods, and Section 6.3 presents and briefly discusses the results. As described in Section 6.4, we improved several aspects of the platform's hardware and software based on user feedback from this study

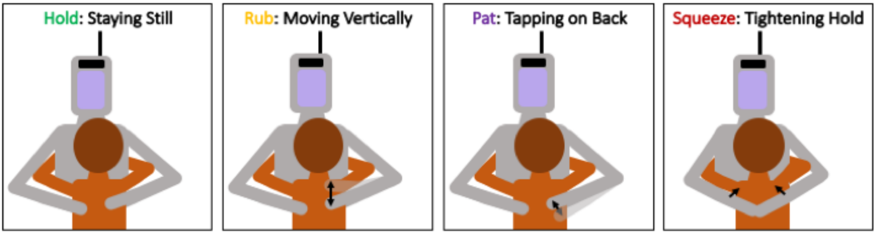


FIGURE 6.1: The four intra-hug gestures that our hugging robot, HuggieBot, can perform, either in response to a user action or proactively when it does not detect any user actions. Despite their importance during prolonged hugs between humans, no prior hugging robot has been able to detect and respond to intra-hug gestures.

as well as the previous chapter's evaluation of HuggieBot 2.0. The collected data were then used to develop a perception system and a behavioral response algorithm for the updated version of our platform, HuggieBot 3.0. Specifically, Section 6.5 explains how we analyzed the microphone and pressure sensor data collected from our novel inflated robot torso (HuggieChest) as 32 diverse users performed four distinct intra-hug gestures. Our developed machine-learning methods quickly detect and reliably classify these different gestures. Based on the 32 users' ratings of the different robot responses, we developed a probabilistic behavior algorithm to determine which action the robot should perform in response to a user gesture; it is described in Section 6.6. Rather than maximizing user acceptance for each robot gesture, which would result in the robot only squeezing the user, our behavior algorithm balances exploration and exploitation (March, 1991) to create a natural, spontaneous robot that provides comforting hugs.

As detailed in Section 6.7, we then ran a follow-up study with 16 new users to test our detection and classification system's real-world accuracy and evaluate the user acceptance of our robot behavior algorithm (validation study). Section 6.8 shares the results of this study, which show that HuggieBot 3.0 is the *first fully autonomous human-sized hugging robot that recognizes and responds to the user's intra-hug gestures.* In creating this new robot, we accept and build upon the six tenets for hugging robots presented in Chapter 5 and Block, Christen, et al. (2021). We agree that a hugging robot should: (T1) be soft, (T2) be warm, (T3) be sized similar to an adult human, (T4) visually perceive and react to an approaching user, (T5) autonomously adapt its embrace to the size and position of the user's body, and (T6) reliably detect and react to a user's desire to be released from a hug regardless of their arm positions. We present a **refined version of T4** plus **five additional tenets** for autonomous hugging robots that we derived from the findings of our two studies:

T4. (refined) A hugging robot should *autonomously invite the user for a hug* when it detects someone in its personal space, and then it should *wait for the user to begin walking toward it* before closing its arms to ensure a consensual and synchronous hugging experience.

T7. A pleasant hugging robot should *perceive the user's height and adapt its arm positions accordingly* to comfortably fit around the user at appropriate body locations.

T8. It is advantageous for a hugging robot to *accurately detect and classify gestures applied to its torso in real time*, regardless of the user's hand placement.

T9. Users like a robot that *responds quickly to their intra-hug gestures*.

T10. To avoid appearing too robotic and to help conceal inevitable errors in gesture perception, a hugging robot should not attempt perfect reciprocation of intra-hug gestures. Rather, the robot should adopt a *gesture response paradigm that blends user preferences with slight variety and spontaneity*.

T11. To evoke user feelings that the robot is alive and caring, the robot should *occasionally provide unprompted, proactive affective social touch* to the user through intra-hug gestures.

Section 6.9 discusses the results of the validation study in the context of these six new hugging tenets, and it also addresses the limitations of our approach. Finally, Section 6.10 provides a summary of this chapter.

6.2 ACTION-RESPONSE ELICITATION STUDY – METHODS

This study serves three main goals. First, we sought additional user comments on all aspects of HuggieBot 2.0 to guide major updates to this platform. Second, this study aimed to collect a large corpus of representative haptic sensor data for four common intra-hug gestures so that we can create a perceptual pipeline that detects and identifies these gestures in real time. Third, this study sought to gather user preferences for how a hugging

FIGURE 6.2: Views of HuggieBot 2.0 (Block, Christen, et al., 2021) ready for a hug and hugging a user. This custom human-sized hugging robot has two padded arms, an inflated torso, and a face screen mounded to a rigid frame. A camera above the screen visually senses the user at the start of the interaction, and torque sensors on the shoulder flexion and elbow flexion joints are used to embrace the user with a comfortable pressure. A microphone and pressure sensor in the back chamber of the torso are used to detect user contact and detect and classify gestures. The user ends the hug by releasing the robot's torso and/or leaning back against the arms.

robot should respond to intra-hug gestures; these opinions will be used to develop a behavior algorithm that enables autonomous hugging robots to respond well to intra-hug gestures that the user performs.

6.2.1 *Robot Platform*

In the previous chapter, Chapter 5, we created and validated a custom hugging robot called HuggieBot 2.0, which we use in this study. As shown in Fig. 6.2, the robot features a v-shaped base that makes it easy for users to get close for a hug. The robot's torso, HuggieChest, is a custom sensing system that simultaneously softens the robot and detects user contacts. It is made of two chambers of PVC Vinyl that are fabricated using a combination of heat sealing and gluing. Inside the chamber located on the back (where users place their hands) is an Adafruit BME680 barometric pressure sensor and an Adafruit electret microphone amplifier MAX4466 with adjustable gain. Both are connected to an Arduino Mega, which sends the data to ROS (Robot Operating System) over serial at about 45 Hz. The pressure sensor is used to detect the start and end of user contact with the back chamber, and both sensors will be used to detect and classify intra-hug gestures in this article.

The robot has two six-degree-of-freedom Kinova JACO arms mounted horizontally on a custom metal frame. These arms were selected for being anthropomorphic, quiet, and safe; their movement speed is limited in firmware, and they can be mechanically overpowered by the user if necessary. Torque sensors at each arm joint are used to automatically adjust to the user's size at the start of the embrace, and torque signals above a threshold give the user a second way to end the hug. The arms are covered in foam pads for softening. The robot's head is a custom 3D-printed case that houses the Dell OptiPlex 7050 minicomputer that controls the entire robot, the Intel RealSense Depth Sensing Camera, the robot's face screen, a small JBL speaker, and many cables. On top of each torso chamber is a Thermophore MaxHeat Deep-Heat Therapy heating pad (35.6 cm $\times$ 68.6 cm). A purple robe and gray sweatshirt are placed on top of the heating pads, and gray mittens cover the robot's end-effectors to create the final robot outfit.

6.2.2 *Ethical Approval, Recruitment, and Participants*

The Ethics Council of the Max Planck Society approved this study under protocol F006B of the Haptic Intelligence Department's framework agreement. The first author recruited participants by email, social media, and paper flyers. All participants were English-speaking volunteers from the local area in Stuttgart, Germany. We ran two users as pilot participants to refine the experimental methods; their data were excluded from analysis because they were not given the same instructions as the later participants. 32 people participated in the study; 12 males and 20 females. The participants ranged in age from 21 to 60 (mean = 30, standard deviation = 7) and came from 13 different countries. Overall, the participants did not have a technical background; many experienced their first interaction with a robot during this study. The 27 participants not employed by the Max Planck Society were compensated at a rate of 8 euros per hour.

6.2.3 *Inclusion/Exclusion Criteria*

All users were required to be over 18 years old. All users were also required to speak and understand English well (roughly at B2 level, though it was not checked). Users who were pregnant or other vulnerable people were excluded from participation. Additionally, no one with uncorrected motor or vision disabilities was allowed to participate. Finally, no one with COVID-19 symptoms or recent exposure to someone with the symptoms or recent visit to a risk zone was able to participate in the study.

6.2.4 *Procedure*

After confirming their eligibility to participate given the exclusion criteria and local COVID-19 regulations, users scheduled an appointment for a 1.5-hour-long session with HuggieBot 2.0. After the participant's arrival, the experimenter explained the protocol using the informed consent document as a guide. The potential participant was given time to read over the consent form thoroughly and ask any questions. If he or she was still willing to participate, the user signed the consent form and the video release form. After receiving both these documents, the experimenter turned on two video cameras to record the experiment.

The user filled out a demographic survey on a tablet computer. Then, the investigator introduced the robot as the personality "HuggieBot" and explained its key features, including the emergency stop. The experimenter explained how the first half of the experiment would work and how the participant could initiate the hug. She also explained the different ways to be released from a hug (release hands from the robot's back or lean back against the robot's arms). At this point, the user filled out an opening survey to document their initial impressions of the robot. We asked users the same questions after the experiment to see how prolonged interaction with the robot affected their responses. The opening and closing survey results were described in the previous chapter, showing significant positive changes in several ratings (Block, Christen, et al., 2021). Users practiced hugging the robot and acclimated to the hug initiation methods and timing of the robot's arm movements before the experiment began. All users then experienced eight hugs that made up the first half of the experiment; as reported in Chapter 5, these results validated the haptic sensing for hug sizing and hug release (Block, Christen, et al., 2021). Interestingly, users showed no preference between starting the hug with a button press or starting it by walking toward the robot (Block, Christen, et al., 2021), an interaction aspect that we seek to improve.

The second half of the experiment contained the activities related to intra-hug gestures. We used two 4×4 balanced Latin squares and a participant number that is a multiple of eight to counter-balance any effects of presentation order (Grant, 1948). Participants experienced a total of sixteen hugs in four groups, each made of four hugs. In each group, the user was instructed to perform a specific action (hold still, rub the robot's back, pat the robot's back, or squeeze the robot) at any point during the hug, as many times as he/she desired. Within a group of hugs, in response to a user action, the robot would perform a different gesture during each hug (staying still, moving vertically, tapping on the user's back, or tightening its hold on the user). When staying still, the robot's arms did not move. For moving vertically (rubbing), the shoulder lift angle of the robot's left arm was increased by 3° and then returned to its original value twice in a row for each rub response. For tapping on the user's back (patting), the elbow flexion joint of the robot's left arm was increased by 3° and then decreased by 6° twice in a row before returning to its original value. For tightening the hold on the user (squeezing), the shoulder flexion joints of both arms were adjusted inward by 1° while both elbow flexion joints were simultaneously adjusted inward by 3°; all four joints were then returned to their original values. Each movement is commanded in joint space relative to the current hug's embracing pose around the user. The joints move between points at the robot's maximum speed, yielding fixed-duration gestures that last approximately 2 seconds.

Because we had not yet developed autonomous intra-hug perception or action capabilities, the timing of each robot response was controlled by the experimenter, who visually observed the actions of the user. A version of Fig. 6.1 was printed on a large poster and placed in the experiment room for the participants to reference at any time. While the robot responses were designed to be the same as what the user was instructed to perform, the poster contained only the pictures and descriptions, without the colored gesture names, to avoid swaying the participants to match robot responses with user actions of the same name. After each hug, the experimenter asked the user which intra-hug gesture they thought the robot had performed; the hug was repeated if the user was not able to identify the robot response correctly, or if the user performed the wrong action.

After a successful action-response hug, users were asked to rate how much they liked that robot response, given the action they performed, using a continuous sliding scale from hate (0) to love (10). To focus on more general principles of human-robot interaction, here we asked users to rate the appropriateness of the gesture response rather than the quality with which HuggieBot 2.0 performed the gesture. After users had experienced all four robot responses for the given user action, they were given time to review and adjust their survey entries before calling the experimenter to explain their ratings. After testing all sixteen hug combinations of user action and robot response, participants were asked to rate the quality of each of the robot responses on the same hate-love scale. After the closing survey, a free-response question asked the user to provide any comments or feedback they had about the experiment.

6.3 ACTION-RESPONSE ELICITATION STUDY – RESULTS

We analyzed the user ratings, pressure sensor and microphone data, and user comments from the action-response elicitation study to understand how HuggieBot 2.0 might be upgraded to become capable of autonomously detecting and responding to intra-hug gestures.

6.3.1 *User Ratings*

As previously mentioned, each user performed each action during four different hugs, experiencing a different robot response during each hug. In total, all users rated sixteen pairs of user actions and robot responses. Figure 6.3 shows the responses to the sixteen different pairs. The color of each cell in the matrix represents the average score from hate (0) to love (10) over all users. The black dots inside each cell show the 32 individual user ratings, always presented in the same order from lowest to highest average user rating across. The three lowest average ratings (5.2, 4.9, and 4.9) all occurred when the user performed an active gesture (rub, pat, or squeeze, respectively) and the robot did not move in response (hold). The hold-hold pairing received a much higher average rating (6.6), as did all conditions wherein the robot responded to user inaction (hold) or action (rub, pat, squeeze) with a rub, pat, and especially a squeeze. While some users gave each action-response pair a rating below neutral, the average ratings achieved for the appropriateness of all of the responsive robotic intra-hug gestures were consistently high.

Figure 6.4 shows the user ratings of the quality of the robot gestures. Users rated the quality of the robot staying still and squeezing as very high

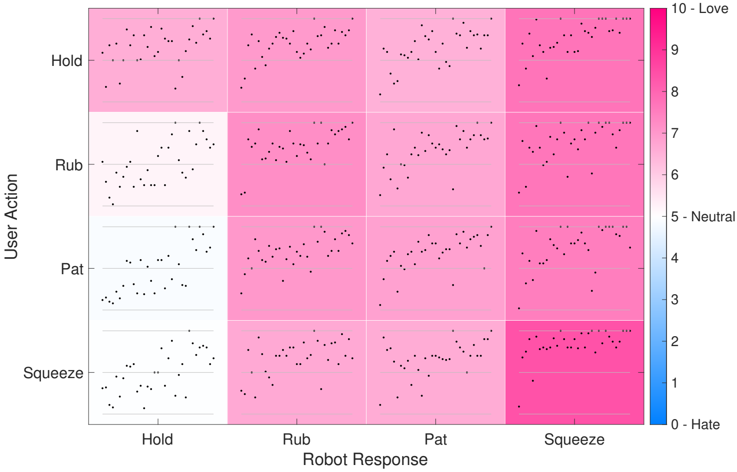


FIGURE 6.3: A matrix showing the user ratings of the appropriateness of each possible robot response to the four intra-hug actions that users performed in the action-response elicitation study. The color of each square represents the average rating, following the legend shown at right. The dots in each cell show the individual rating of each user, consistently ordered based on their average score from low to high. The pale horizontal lines in each square show the ratings of 0 (hate), 5 (neutral), and 10 (love).

(8.5 and 8.2, respectively); no user gave a negative rating for the robot's hold, and only two of the 32 users gave negative ratings for squeeze. Rubbing and patting were rated positively (5.8 and 5.7, respectively), but closer to neutral. The fact that these gesture quality ratings differ somewhat from the average response appropriateness ratings in Fig. 6.3 shows that users were at least moderately successful at distinguishing these rating tasks from one another, particularly regarding hold.

6.3.2 *Pressure Sensor and Microphone Data*

Data were recorded from HuggieChest's pressure sensor and microphone for each of the 16 hugs for all 32 users, yielding a total of 512 recordings. Analyzing these signals after the experiment shows common characteristics between different users performing the same gesture. As a representative

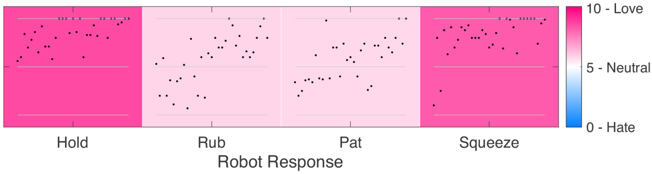


FIGURE 6.4: A matrix showing the user ratings of the quality of the four robot responses, using the same visualization approach as Fig. 6.3.

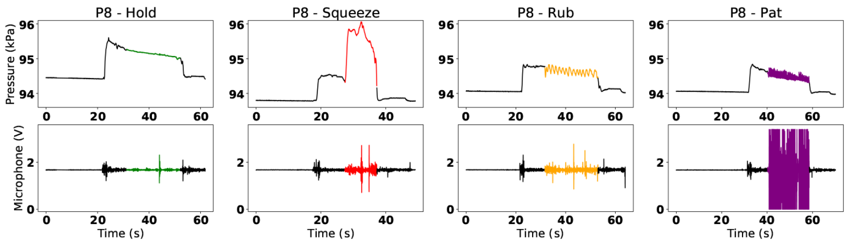


FIGURE 6.5: Sample pressure signals and microphone signals from participant 8 performing each of the four gestures during the action-response elicitation study. The colored data points mark the time periods manually labeled as positive examples of the indicated intra-hug gestures.

sample, Figures 6.5 and 6.6 show the pressure and microphone data collected from two different participants performing the same gestures. While the torso chamber was inflated to somewhat different initial inflation levels for each participant, characteristic signals can still be recognized to determine the type of contact made, regardless of inflation. The use of these two sensors allows us to differentiate between both coarse and fine contacts. While the pressure signals look similar between rubs and pats, we can differentiate between the two by the pat's much larger microphone response. Additionally, the microphone signals look quite similar for squeezes and rubs, but looking at the corresponding pressure signals allows us to determine which action is being performed; squeezing the robot drastically increases the chamber's pressure, while rubbing it does not. These results show the benefit of using two sensors for detecting intra-hug gestures.

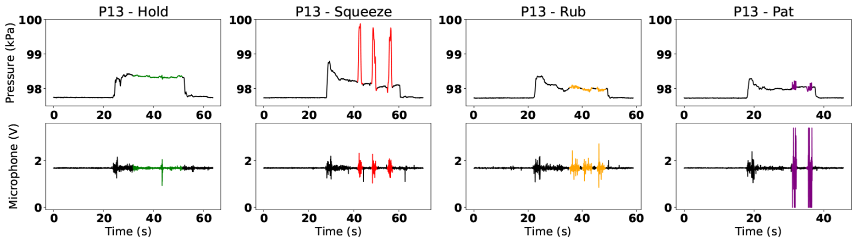


FIGURE 6.6: Sample pressure signals and microphone signals from participant 13 performing each of the four gestures during the action-response elicitation study. The colored data points mark the labeled data segments.

Video review revealed that these two participants had different approaches to performing the gestures, which can be recognized in the recorded data. P8 (Fig. 6.5) squeezed the robot only once and performed a continuous strong patting movement. In contrast, P13 (Fig. 6.6) performed three distinct squeezes and two pats, separating each repeated gesture with a short pause (hold). The variety of ways in which the 32 users performed these four gestures was surprising and underscored the importance of gathering a large corpus of sensor data.

Finally, although we asked users to only perform a single type of gesture per hug, we noticed that seven of the 32 participants (21.9%) accidentally combined gestures and sometimes performed two gestures at once. This particularly happened during rubs and pats, when the user would sometimes also unintentionally squeeze the robot.

6.3.3 *User Comments*

Our users' written and spoken comments provide crucial information on how to improve the quality of the hugs HuggieBot can deliver. A systematic analysis reveals several key themes repeated by many users. The majority of users (68.75%) commented that they preferred *not* having to press a button to initiate the hug. The slow speed of the robot's arms and thus the amount of time that it took for the arms to close around the user detracted from the experience (34.37%), particularly for hugs initiated when the user walked toward the robot. Next, almost half of the users (43.75%) mentioned that they could not feel both arms fully against their backs in at least one hug. Some users (21.87%) also commented that the robot's hand placement was inappropriate – either too high on the back (close to the neck) or too low (on the buttocks), both of which made them uncomfortable. Incomplete arm closure and too high or too low hand placement made it difficult for some users (21.87%) to feel the robot performing gestures on their back, which most likely contributed to the variety of user ratings reported. These comments show that improvements are needed for how HuggieBot initiates hugs and adapts to its hugging partner.

Participants in this study experienced the robot both responding to their gestures and holding still (not responding) when they performed a gesture. Almost all users (78.13%) commented that having the the robot respond to their gestures made it feel more "alive," "social," and/or "realistic". As we initially expected users to prefer a robot that reciprocates their gestures, we were surprised to find users enjoyed variety in the responses. When they explained their response ratings to the experimenter, twenty of our users (62.5%) mentioned that reciprocation of their actions every time felt "too mechanical." Rather than thinking the robot made a mistake when it performed a different gesture, users appreciated gestures of similar "emotional investment levels" (P21) and felt it showed "the robot understands [them] and makes his own decision" (P30).

In agreement with the high ratings shown in Fig. 6.3, many users shared positive comments about being squeezed by the robot, saying phrases such as "I love warm, tight hugs" (P7, P8, P17, P21, P25) or that being squeezed by the robot felt "the most natural" (P15) and "the closest to a real human hug, the best response" (P2, P16, P20). Some even went as far as to say that the squeezes gave them "a sense of security and comfort" (P9, P20, P23). In general, users thought the duration of the robot's timed squeeze was "too short" (P5, P24). Additionally, several users suggested making the robot's squeeze duration match theirs because the fixed timing felt "too mechanical" (P1, P6) or like "the robot wasn't as emotionally invested as [they] were" (P21, P28). These comments hint at the need to treat modal gestures such as squeeze differently from event-based gestures such as rub or pat.

Finally, another unexpected finding that wove throughout the comments was how strongly the users anthropomorphized the robot. The experimenter always called the robot "it" or "HuggieBot," yet in both written and verbal comments, 96.87% of users referred to the robot as a "he" when describing how the hugs felt, often explaining a social situation it reminded them of. Such interactions included "a comforting hug from a mother" or "a distant relative at a funeral," "seeing friends at a football match," "receiving a pity hug from someone who doesn't want to," "hugging an ex," to even "hugging a lover." They attributed emotions, mood swings, and attitudes to the robot, depending on how well it hugged them in each trial of the study.

6.3.4 *Brief Discussion*

The ratings gathered in this study provided essential information about user preferences for how a hugging robot should perform intra-hug gestures. We conducted this experiment to guide hardware and software improvements to HuggieBot 2.0, as will be discussed in the following Sections. However, the results can also be generalized and applied to other hugging robots, including those that do not have any haptic sensing. As can be seen in Fig. 6.3, regardless of the user action, a squeeze was always perceived on average as the most enjoyable robot response, including when the user had not just actively performed an action (after user hold). Therefore, researchers who want to improve user opinions of their hugging robot without investing in haptic sensing capabilities should program their robot to occasionally squeeze the user. Simultaneously, the neutral average reactions users showed when the robot did not respond to their rubs, pats, and squeezes indicate that perceiving and responding to intra-hug gestures could greatly improve hugging robots.

6.4 IMPROVEMENTS TO THE PLATFORM

The results of the previous chapter featuring HuggieBot 2.0 (Chapter 5) and the user comments from the action-response elicitation study (Section 6.3.3) showed four main aspects of the system that could benefit from improvement: the hug initiation process, the vertical placement of the robot's arms on the user's body, the reliability of the inflated torso, and the quality and consistency of the robot's embrace around the user's body. Thus, we spent time addressing these concerns to improve the quality of the hug that this robot can deliver to users, upgrading HuggieBot from version 2.0 to version 3.0. Table 6.1 summarizes the key features of these two successive versions. We extensively piloted all of these changes with representative users and made further adjustments based on their feedback. The following subsections provide more detail about the final changes made and used in the validation study.

	HuggieBot 2.0 (Block, Christen, et al., 2021)	HuggieBot 3.0 (new)
Hug Initiation	Single-stage initiation: the experimenter tells the user when the robot is ready. HuggieBot initiates a hug either when the user presses a key or when it visually detects that a user is walking toward it. The robot speaks after the hug has been initiated. The user must adjust his/her walking speed to match HuggieBot's movements and wait for the arms to close.	Consensual two-stage initiation: After detecting a user in its personal space, HuggieBot opens its arms and verbally invites the user for a hug. HuggieBot waits for the user to start walking toward the robot before starting to close its arms.
Arm placement	Constant arm lift angles suitable for an adult of average height.	Automatic adaptation of arm lift angles to the user's observed height (equations (6.1) and (6.2)).
Torso	Two warm inflatable chambers with a smaller chamber in the back compared to the front. Each chamber has a microphone and a pressure sensor, with wires exiting at the bottoms of the chambers.	Two warm inflatable airtight chambers of equal size on the front and back. Only the back chamber has a microphone and a pressure sensor, and the wires exit from the top of the chamber.
Embrace	Automatic adaptation of arm closure to the user's size by thresholding individual joint torque measurements on the four involved arm joints. Wrist flexion joints oriented horizontally, which causes local high pressure on the user's back at the points of contact with the wrist.	Automatic adaptation of arm closure to the user's size with better robustness and comfort through filtering of torque values, adjusted torque thresholds, and adjusted final arm closure angles for the four involved arm joints. Wrist flexion joints oriented vertically so that the surfaces in contact with the user's back are flat.
Hug release	Dual activation: the user removes his/her hands from the robot's back (pressure threshold) and/or leans back against the robot's arms (torque threshold).	Dual activation: same as HuggieBot 2.0 with a reduced torque threshold to allow for easier release.
Gesture perception	Wizard of Oz (by a human operator)	Autonomous: random forest algorithm based on one second of data from the microphone and pressure sensors (Section 6.5).
Behavioral response	Wizard of Oz (by a human operator)	Autonomous: probabilistic response based on mean user preference ratings with slight variety (Fig. 6.3, equation (6.3)), including some pro-active gestures during periods when no user gestures are detected (Section 6.6).

Table 6.1: Summary of HuggieBot 2.0 and 3.0 features.

6.4.1 *Hug Initiation*

The initial evaluation of HuggieBot 2.0 in the previous chapter, Chapter 5, tested two different ways for users to initiate a hug, always starting about 2.5 meters in front of the robot. In the first method, the user pushed a button to start the hugging process. The second method used HuggieBot's built-in depth camera to recognize a human and then start the hugging process when the potential user starts walking toward the robot. Users did not rate the two methods significantly different; this indifference can be explained by their comments on this topic. Because the visual hug-initiation method was triggered by the user's forward movement, and because the robot arms close slowly, users would often reach the robot before its arms had closed very far, causing them to have to wait for the robot's embrace.

After piloting several alternative approaches, we improved the visual hug initiation process by dividing it into two steps. First, when all the necessary software nodes are running, the robot lifts its arms and asks the user "Can I have a hug, please?" The phrase and arm movement clearly show the user when they may begin hugging the robot; previously, the experimenter prompted users when they could begin. Lifting the arms also beneficially reduces the distance the robot's arms need to travel to close around the person, thus shortening the waiting time disliked by users. After this invitation step, HuggieBot uses the previous method to visually detect the user's forward motion and initiate the closing sequence. The robot waits between the "hug request" pose and closing its arms for as long as the user needs in order to reduce time pressure, so users do not feel like they have to start the next hug immediately. These small changes were implemented to make the robot's hug timing more natural and intuitive. This method also beneficially reduces experimenter interaction with the user and better mimics human-human hugs, where one person lifts their arms for a hug and waits for their partner to approach before wrapping their arms around them.

6.4.2 *Adjustment to User Height*

As shown in Fig. 6.4, the robot's rubs and pats received lower average quality ratings than hold and squeeze. When users explained the low ratings, the most common criticism was the location at which the gesture was performed, which was not optimal for their bodies. As HuggieBot 2.0's arms always hugged at the same height off the ground, these gestures were performed too low for tall users and too high for short users. In addition to the inappropriateness of some contact locations reported in the comments, the convex or concave curvature of different areas of different users' backs exacerbated this problem by causing loss of contact or excessive contact when the robot performed some gestures.

To resolve this hand placement issue, the robot must improve its visual perception of the user. In addition to detecting a potential user and estimating their approach speed toward the robot, HuggieBot needs to perceive the user's approximate height and adjust its arm positions accordingly, something humans do naturally, quickly, and efficiently.

To simplify this problem, several assumptions were made. First, we assumed that the camera is perfectly parallel to the floor. Second, we assumed that the person approaching is standing perpendicular to the floor. These assumptions help simplify the problem from a three-dimensional problem in point-cloud space to a planar problem. Simplicity is desired in this case to keep the computational load low and allow for real-time adjustments. The problem then becomes one of similar triangles.The depth camera's resolution is 1280 pixels $\times$ 720 pixels, and its focal length is 651.55 pixels. Based on the room's size constraints and the need to keep the camera oriented parallel to the floor, the camera cannot see the user's feet and lower legs when the person is first detected. To accommodate this reduction in the bounding box's size, once visible height of the user has been calculated in meters, a small adjustment must be added based on the user's distance from the camera to account for the height of the unseen portion of their body. The full linear projection can be written as follows, using constants obtained through measurements:

$$H = \frac{(D \cdot b)}{f} - 0.5518 \cdot D + 1.73 \text{ m} \tag{6.1}$$

where H is the user's full height in meters, D is the distance between the user and the robot in meters, f is the depth image focal length in pixels, and b is the height of the bounding box of the object detection of a person in pixels. We found that individual height estimates computed in this way are somewhat noisy, so HuggieBot 3.0 averages five successive measurements.

We set the ideal shoulder lift angle for HuggieBot 3.0's left arm based on the estimated height of the user, and we then offset the right arm shoulder lift angle up by 20 degrees from that point to create good inter-arm spacing. To determine appropriate arm lift angles for users of different heights, we performed brief experiments with two model users at the minimum (1.40 m) and maximum (1.93 m) user heights we anticipate encountering. We manually adjusted both robot arms around each model user to a comfortable height on their back and recorded the corresponding shoulder lift joint angles. We perform linear interpolation to find the ideal left shoulder lift joint angle for the approaching user, as follows:

$$\theta_\ell = \theta_{\ell,\min} + (H - H_{\min}) \cdot \frac{(\theta_{\ell,\max} - \theta_{\ell,\min})}{(H_{\max} - H_{\min})} \quad (6.2)$$

where $\theta_{\ell,\min}$ and $\theta_{\ell,\max}$ are the robot's left shoulder lift angle angle for the minimum-height and maximum-height model users, respectively, H is the user's estimated height in meters, and $H_{\min}$ and $H_{\max}$ are the height of the short and tall model users, respectively. When the user's estimated height is outside the range of the model users, the closer model user's robot arm placement is used.

6.4.3 *New Torso*

We created a new inflated torso for HuggieBot 3.0 to address several shortcomings in the previous design. HuggieBot 2.0's torso contains a pressure sensor and a microphone in both the front chamber and the back chamber. Early testing showed that the front chamber data provided little information beyond the back chamber data, so HuggieBot 2.0 does not use the information from these sensors. Our new torso has sensors only in the back chamber; furthermore, the sensor wires exit the top of the chamber rather than the bottom of the chamber to minimize the distance to the computer in the robot's head.

The torso's initial design featured two different-sized chambers, with the back chamber being slightly smaller. Study participants of all sizes used various arm positions to hug the robot and perform intra-hug gestures on its back. Therefore, we decided to increase the back chamber's size to be equal to the front chamber to better accommodate all users.

Some users squeezed the robot much more tightly than we anticipated during the squeezing hugs of the action-response elicitation study, occasionally popping holes in a chamber or forcing a resealable inflation valve to open; both of these failure modes allow air to escape, change the feel of the robot's torso, and require re-inflation. We designed the new chamber to be more robust to withstand these higher pressures. We ensured a robust and airtight seal on HuggieBot 3.0's new torso by heat sealing along the edges and then using HH-66 Vinyl Cement on top of the heat seal.

The newly constructed torso was tested by pilot users performing the four studied gestures during hugs with the robot. After we matched the sensitivity of the new microphone to that of the previous one, the sensor recordings exhibited the same general patterns as those shown in Figs. 6.5 and 6.6. The lack of leaks in the new chamber beneficially also stabilized the starting pressure, which makes both the feel of the robot and the measurements of its haptic sensors more consistent over time.

6.4.4 *Quality and Consistency of the Embrace*

Finally, several smaller changes were made to how the robot's arms grasp the user based on the feedback from the action-response elicitation study as well as additional pilot users. The changes we made are as follows:

1. Some users were thinner than we had anticipated, so some of the robot's joints reached their goal angles without coming into contact with the user's back. Therefore, we made the closing goal angles much smaller, such that no user will fit inside the goal pose, which ensures good contact between both robot arms and the user's back.

2. To automatically adjust to our users' shapes and sizes, we use a torque threshold to turn off individual joints when they come into contact with the user. If a single torque measurement exceeds this threshold, HuggieBot 2.0 stops that joint from moving further. We found that while this kept users safe, it did not provide them with a consistent feeling of being fully embraced because some spuriously high torque readings were occasionally measured. Therefore, we implemented a moving average filter on all torque measurements with a window of three values. When the average torque in this window surpasses the threshold, the joint stops. After making this adjustment, we also tuned the torque thresholds and the relative offset for each joint's final target with multiple diverse users to ensure HuggieBot 3.0's embrace was comfortable and not painful.

3. With the improved arm placement and more complete closure, we noticed that the wrist rotation of the robot's arms could cause uncomfortably high pressure on the backs of some users. We adjusted the goal pose for the wrist angles for both arms to ensure the flat and comfortable side of the wrist is in contact with the user's back.

4. With the improved quality of the contact between the arms and the user's body, we found that HuggieBot's squeeze was now too tight

for many users. We thus reduced the elbow joint's squeeze movement from a magnitude of 5° to only 3°.

5. We reduced the torque threshold required to release from the hug during gestures. This value was set rather high (60 Nm) for HuggieBot to prevent users from accidentally triggering a torque release when performing a gesture; however, pilot users were occasionally bothered by how hard they had to lean back to make the robot release them in the middle of a gesture. This threshold was tuned through pilot testing to have a final value of 20 Nm during a regular embrace and 40 Nm during robot gestures. It is important to note that we kept a single value for this torque threshold rather than checking a moving average of the measurements in order to release the user without delay. If a user wants to end the hug at any time, HuggieBot 3.0 opens its arms immediately.

These four categories of hardware and software improvements upgraded HuggieBot 2.0 to become HuggieBot 3.0, as summarized in Table 6.1.

6.5 DETECTION AND CLASSIFICATION OF GESTURES

To build an autonomous perception system that HuggieBot 3.0 can use to detect and classify intra-hug gestures, we manually labeled 498 trials (14 of the initial 512 trials were discarded due to problems in the data) recorded in the action-response elicitation study. One of the authors visualized the pressure and microphone signals and marked the hug's start to be the first point where the pressure rises from its initial value. Similarly, the hug's end was the first point where the pressure value declined back to its initial value. The same author marked the start and end of touch gestures based on the distinct signatures of the gestures in the pressure and microphone signals.

Next, we divided the time-series hug data into a large number of segments and extracted statistical features from each segment. We applied a moving window (W) with an overlap size (O) to divide the time-series data for each hug into shorter segments. If a segment's overlap with the gesture timestamps was above a certain threshold (T), the segment was labeled with that touch gesture. For each segment, we subtracted the baseline pressure and baseline microphone values from the corresponding signals. These baseline values were calculated by taking the median of the first 150 pressure and microphone data points after the start of the hug, which corresponds to about 3.3 seconds of data recorded at 45 Hz. We extracted 80 statistical features for each segment, including sum, minimum, maximum, average, median, standard deviation, variance, number of peaks, interquartile range, and area under the curve for the pressure and microphone signals and their first and second time derivatives.

We divided the segments into train, validation, and test sets. We trained a random forest algorithm on 70% of the data and used 20% of the data as the validation set to determine the impact of the three parameters W, O, and T on the model performance. We tested window sizes of W = 50, 75, and 100 data points, overlap sizes of O = 37, 25, and 12 data points, and labeling thresholds of T = 0.75, 0.5, 0.25. The classification accuracy varied by only about 1% to 3% among different parameter combinations. However, larger values of W and T would make the robot slower at detecting user gestures. Thus, we selected W = 50, O = 37, and T = 0.75 as a trade-off between gesture detection accuracy and delay. Fig. 6.7 shows the confusion matrix after applying the resulting random forest model to the remaining 10% test data. The overall classification accuracy achieved is 88%, with the best performance on detecting squeezes.

After development and offline validation, this perceptual pipeline was transferred to the physical robot and adapted to run in real time. To conserve computational effort while achieving a fast reaction time, we calculate features on the most recent window of 50 data points every 10 samples, rather than every sample. Intra-hug gestures are thus detected at a rate of approximately 4.5 Hz, always yielding an output of hold, pat, rub, or squeeze. After being trained on the dataset collected from HuggieBot 2.0, the classifier was tested on the new inflated torso (Sec. 6.4.3) of HuggieBot 3.0. Pilot testing showed that the classifier's performance was worse than desired, probably due to the changes in the size and shape of the back chamber. Gesture samples were thus acquired from five pilot participants performing each of the four studied gestures during two hugs on the new torso. These 40 trials were labeled and processed in the same way as the original data from the action-response elicitation study. Because these additional examples did not have sufficient power when added to the 70% training data (354 trials), we trained a new version of the classifier using 80% of the data from 10 randomly selected users from the action-response study (125 trials) plus the newly collected 40 trials. This retrained version showed good performance on the new inflated torso in pilot testing, so it was selected for use in the validation study.

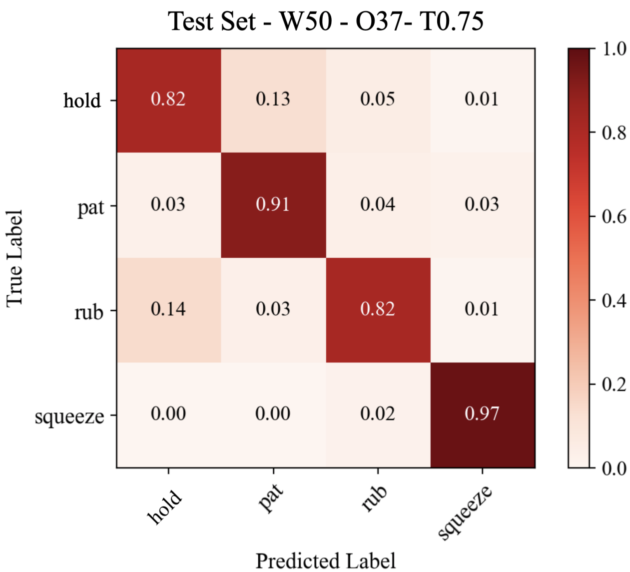


FIGURE 6.7: Confusion matrix for the test dataset gathered in the action-response elicitation study.

6.6 BEHAVIORAL RESPONSE TO DETECTED GESTURES

After it possesses a good pipeline for detecting and classifying intra-hug gestures, a hugging robot needs to decide how to act, i.e., which of the available intra-hug gestures to perform. We implemented the robot's behavior in all situations using a simple probabilistic approach (Section 6.6.1), calculating the likelihood that the robot will perform each gesture as a function of the relevant average user ratings given the action that was just detected. Although we have treated them equivalently thus far, the active gestures of pat, rub, and squeeze occur far less often than the passive gesture of hold in natural hugs. The hold gesture can be thought of as the standard background against which the active gestures occur; almost every hug in the action-response elicitation study included periods of time where the user was simply holding the robot. The differences between active and passive gestures require somewhat different types of behavioral responses from the robot. We found the need for an additional distinction between the discrete active gestures of pat and rub (Section 6.6.2), which occur in discrete units, and the modal active gesture of squeeze (Section 6.6.3), which involves transitioning into and out of the state of applying higher pressure to the hugging partner. Note that the described methods are not limited to these four intra-hug gestures: another example of a discrete active gestures that future hugging robots could consider is tickling, while an alternative modal active gesture could be leaning into one's hugging partner. Finally, we describe how to apply our probabilistic behavior paradigm when a hugging robot detects a passive gesture such as hold (Section 6.6.4). Figure 6.8 depicts the robot's state transition diagram to clarify the robot's controller.

6.6.1 *Probabilistic Behavior Paradigm*

Examination of Fig. 6.3 shows that each row follows a somewhat different distribution of average user ratings; thus, the appropriateness of a particular robot response critically depends on the action the user has just performed. Our probabilistic approach was designed to respect that dependency as well as the relative appropriateness of the different favorably received responses in the way it chooses between gestures.

Specifically, we designed a simple generic method for converting the user preferences gathered during the action-response elicitation study into a probabilistic behavior model that determines which action the robot should perform based on which gesture it just detected from the user. The equation we crafted to transform the average ratings into probabilities is as follows:

$$p_{g|a} = \frac{(r_{g|a} - \eta)^m}{\sum_{i=1}^{N} (r_{i|a} - \eta)^m} \tag{6.3}$$

where $p_{g|a}$ is the probability with which the robot will perform the specified gesture g given the user action a, $r_{g|a}$ is the average user rating that gesture g received when presented as a robot response to user action a, η is the neutral value on the rating scale, m is a positive power that controls how strongly higher-rated gestures should be favored (chosen to be 3 for our study), N is the total number of gesture options being considered (usually 4 for our study), and $r_{i|a}$ is the average user rating that gesture i received when presented as a robot response to user action a. This formula subtracts the neutral value $\eta = 5$ from each average rating to focus only on responses that were positively received; if a rating is below neutral, the result of this subtraction is set to zero to ensure the robot never performs that gesture in response to this user action.

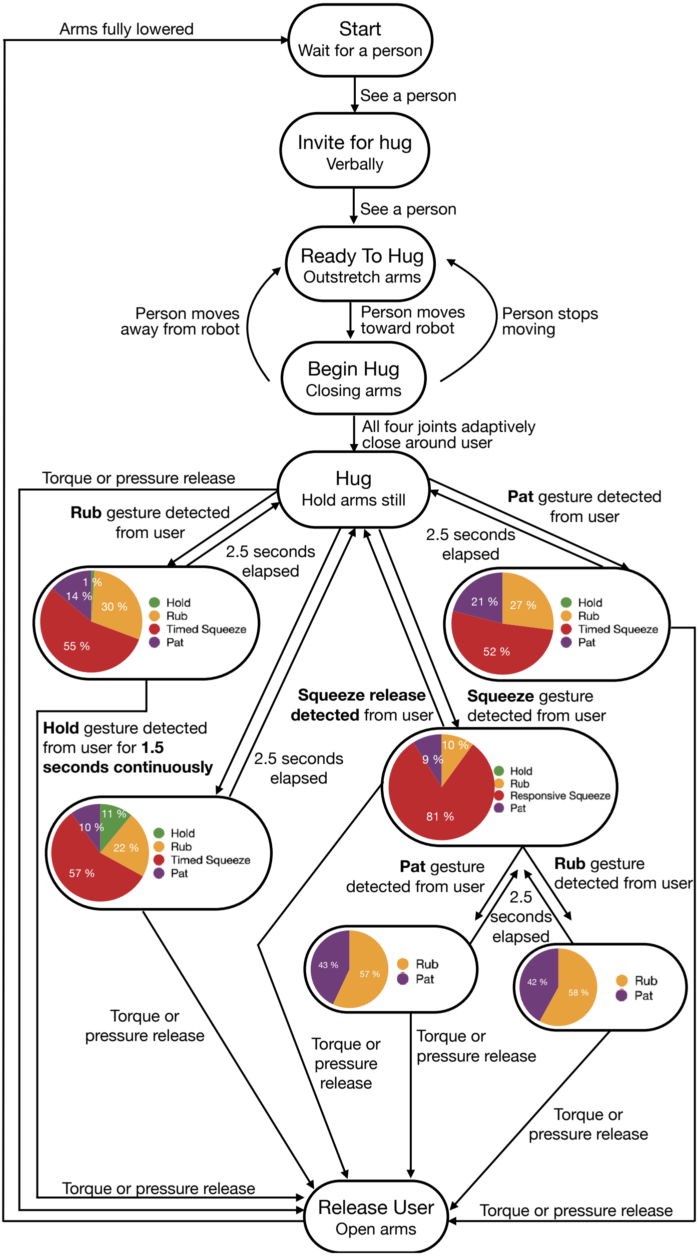


FIGURE 6.8: Robot state transition diagram.

The numerator and denominator terms in equation (6.3) are raised to the power of $m = 3$ to increase the probability the robot will select highly rated responses; other powers could be used to provide other blends between exploration of different options and exploitation of the known best choice. For example, $m = 1$ sets the gesture probabilities to be directly proportional to the relative strength of their received ratings, which we found to be too random. In comparison, $m = 0$ would perform the N gestures with equal probability (pure exploration with no dependence on user rating), and $m = \infty$ would always select the highest-rated gesture (pure exploitation with no variety). Regardless of the value of m, the N resulting probabilities necessarily sum to unity.

In practice, we pre-compute the contingent gesture probabilities from the average ratings shown in Fig. 6.3. The robot's software then implements this probabilistic behavior algorithm by generating a random number between 0 and 1 each time a response is required. The probabilities relevant to the detected action are stacked, and HuggieBot 3.0 executes the gesture corresponding to the generated number.

6.6.2 *Responding to Discrete Active Gestures*

The intra-hug gestures of rubbing and patting both consist of discrete hand motions that the user can perform a single time or repeat many times in a row. Thus, the robot's response to these gestures is designed to also follow a discrete action paradigm. When the perceptual pipeline detects the start of a new user rub or pat, the robot first determines which discrete gesture to respond with by applying equation (6.3) to the average user ratings from the appropriate row (rub or pat) of Fig. 6.3. Because hold received a positive average rating in response to user rub actions, the robot chooses between all four gestures in this situation with probabilities of $p_{\text{hold}\,|\,\text{rub}} = 0.01$, $p_{\text{rub}\,|\,\text{rub}} = 0.30$, $p_{\text{pat}\,|\,\text{rub}} = 0.14$, and $p_{\text{squeeze}\,|\,\text{rub}} = 0.55$. In contrast, hold received a slightly negative average rating in response to user pat actions, so the hold response is never chosen after a pat is detected ($p_{\text{hold}\,|\,\text{pat}} = 0.00$); the robot's three non-zero gesture response probabilities in this case are $p_{\text{rub}\,|\,\text{pat}} = 0.27$, $p_{\text{pat}\,|\,\text{pat}} = 0.21$, and $p_{\text{squeeze}\,|\,\text{pat}} = 0.52$. The gestures with which the robot responds are all executed with fixed timing, as in the action-response elicitation study.

The overlapping time windows used in our perception algorithm cause many successive detection events to be triggered over the course of a single long user action. Thus, while the robot is executing the response to a detected discrete active gesture, it enters a state of ignoring new detected actions for 2 seconds (roughly the time it takes the robot to perform the selected gesture). This way, the robot does not accumulate a backlog of queued actions it needs to respond to, which could result in a never-ending pat or rub response. It should be noted that even in this state, the robot constantly checks to see if the user has indicated a desire to end the hug by either releasing pressure on the back chamber or leaning back against the arms. If either of these conditions occurs, the state-machine overrides whatever gesture the robot is performing and begins releasing the user. Otherwise, the robot returns to responding to new detected actions after the two-second-long delay elapses.

6.6.3 *Responding to Modal Active Gestures*

When a new squeeze is detected, HuggieBot 3.0 calculates the response gesture probabilities by substituting the average user ratings from the squeeze row of Fig. 6.3 into calculating equation (6.3). Because hold received a slightly negative average rating in response to user squeezes, the hold response is never chosen after a squeeze is detected ($p_{\text{hold}\,|\,\text{squeeze}} = 0.00$); the robot's three non-zero gesture response probabilities in this case are $p_{\text{rub}\,|\,\text{squeeze}} = 0.10$, $p_{\text{pat}\,|\,\text{squeeze}} = 0.09$, and $p_{\text{squeeze}\,|\,\text{squeeze}} = 0.81$. If the chosen gesture is a discrete rub or pat, the response proceeds as described in the previous Section. However, the squeeze response to a squeeze action merits special handling because of its fundamentally different nature and popularity of this robot gesture among users.

Unlike the repeated discrete motions of rubs and pats, squeezes are modal active gestures that move between two states that apply lower and higher pressure on the partner, respectively. As shown by the action-response elicitation study, user squeeze actions can vary greatly in duration, and the start and end of a robot squeeze are quite salient to the user. To avoid repeatedly squeezing the user for fixed time intervals, and to follow the user suggestion to have the robot's extended squeeze duration match theirs, HuggieBot 3.0 responds somewhat differently to squeezes compared to discrete active gestures. Specifically, if the robot's response to a detected squeeze is to squeeze the user back, as soon as the robot detects a squeeze, it enters a squeezing state, and it leaves this state only when the perception algorithm detects a hold (which means the chamber pressure has decreased to near the baseline value and that the microphone is not detecting strong activity). At this point, the robot stops squeezing and returns to the normal hug state, where it reacts to detected gestures in the normal way.

The robot's response to user squeezes has one other unique behavioral capability. Observing user behavior during the action-response elicitation study showed that several users squeezed the robot at the same time that they performed rubs or pats. Thus, when the robot is in the squeezing state, it continues monitoring the detected gestures. It stays in the squeezing state when the perceptual algorithm continues detecting user squeezes. If a rub or a pat is instead detected, it responds to this discrete active gesture layered on top of a modal active gesture in a way similar to usual discrete active gestures. The robot's response is again determined using the probabilistic behavior paradigm given in equation (6.3). Because the robot is already in the squeezing state, the robot squeeze response is not considered, so the robot chooses between the remaining response options for the detected action, which are all discrete and thus take a fixed amount of time. In this way, a user can squeeze the robot, be squeezed in response, then rub or pat the robot's back, and receive a rub or a pat in response, still during the squeeze.

6.6.4 *Responding to Passive Gestures*

The final case happens when the perceptual pipeline detects that the user is passively holding the robot and not performing an active gesture. Based on the positive user ratings and comments found for proactive robot gestures in the action-response study, it is important for hugging robots to occasionally take the initiative to perform a gesture when it detects only hold gestures from the user. Providing proactive affective touch is more subtle than responding to discrete actions because the robot must take the temporal element into account to ensure that the person is genuinely not doing anything over an extended period. To determine the frequency at which proactive robot affective touch should occur, we looked at the rate at which the investigator remotely activated the gestures in the action-response elicitation study. The average time delay between the end of one proactive robot gesture and the beginning of the next one across the holding hugs of the 32 participants was 1.5 seconds. Thus, the behavior algorithm waits until it detects the hold user action detection for 1.5 seconds in a row (approximately 7 overlapping windows in a row) before proactively initiating an intra-hug gesture by the robot. Our probabilistic behavior model is again used to determine which gesture the robot should perform, as described in equation (6.3); the calculated conditional probabilities are $p_{\text{hold}\,|\,\text{hold}} = 0.11$, $p_{\text{rub}\,|\,\text{hold}} = 0.22$, $p_{\text{pat}\,|\,\text{hold}} = 0.10$, and $p_{\text{squeeze}\,|\,\text{hold}} = 0.57$. This definition completes our description of how HuggieBot 3.0 behaviorally responds to gestures detected by its perceptual pipeline.

6.7 VALIDATION STUDY – METHODS

This study aims to test and validate the new platform of HuggieBot 3.0 (Section 6.4), its perceptual pipeline for detecting and classifying intra-hug gestures (Section 6.5), and its probabilistic behavior algorithm for responding to detected gestures (Section 6.6). Regarding the platform improvements, we specifically sought to evaluate the user response to the updated hug initiation process (Section 6.4.1) and the updated robot hand placement based on estimated user height (Section 6.4.1). The validation study we conducted was similar to the action-response elicitation study except for the key difference that the robot was always behaving autonomously. The Max Planck Society's Ethics Council approved all methods for this study under a new framework agreement protocol number, F011B. The investigator recruited voluntary participants in the same manner as described in Section 6.2.

6.7.1 *Participants*

The sixteen participants were English-speaking volunteers recruited from Stuttgart, Germany. No participants were employed by the Max Planck Society, so all were compensated at a rate of 8 euros per hour. This study was every user's first time interacting with any version of HuggieBot and the first robotic interaction of any kind for most users. Half of our participants were men, and half were women; 93.75% were non-technical. The average height reported by our users was 1.69 m, with a standard deviation of 0.10 m. The participants ranged in age from 22 to 38 (mean = 30, standard deviation = 4.76) and came from 10 different countries. Over 80% of the participants identified as enjoying receiving hugs from others.

6.7.2 *Inclusion/Exclusion Criteria*

All users were required to be over 18 years old. All users were also required to speak and understand English well (roughly at B2 level, though it was not checked). Users who were pregnant or other vulnerable people were

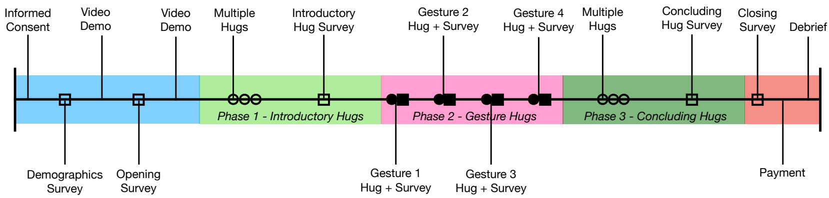


FIGURE 6.9: The timeline of the validation study. The colored boxes represent the five phases of the experiment. The left-most vertical line marks the start of the user study, and the right-most vertical line marks the end. Vertical lines coming off the center timeline show the order of the user's activities. Squares on the timeline indicate when the user filled out a survey, while circles indicate when they performed hugs. If the shape (square or circle) is not filled, that activity was performed in an open-ended way. If it is filled in with black, the activity was more controlled. Note that each user chose the number of open-ended hugs to perform in phases 1 and 3 rather than always performing the three hugs indicated in the timeline.

excluded from participation. Additionally, no one with uncorrected motor or vision disabilities was allowed to participate. Finally, no one with COVID-19 symptoms or recent exposure to someone with the symptoms or recent visit to a risk zone was able to participate in the study.

6.7.3 *Procedure*

Like the action-response elicitation study, prospective users were required to confirm their eligibility to participate by email, given the exclusion criteria and the local COVID-19 regulations. The timeline for this study can be found in Fig. 6.9. Once the recruit arrived at the experiment site, the investigator explained the study. They read over and signed the informed consent document and video release form after asking the experimenter any questions. At this point, the experimenters started recording the study and preparing the robot while the user filled out the demographics survey.

Instead of verbally explaining how to interact with HuggieBot 3.0, the investigator showed the user a video on a laptop of a sample user hugging the robot; the video demo is included as supplementary material for this article. Without providing any verbal instructions, this video shows the user how the robot asks for a hug (lifting its arms and saying "Can I have a hug, please?"). It also shows how to prompt the robot to close its

Question
I feel understood by the robot
I trust the robot
Robots would be nice to hug
I like the presence of the robot
I think using the robot is a good idea
I am afraid to break something while using the robot
People would be impressed if I had such a robot
I could cooperate with the robot
I think the robot is easy to use
I could do activities with this robot
I feel threatened by the robot
This robot would be useful for me
This robot could help me
This robot could support me
I consider this robot to be a social agent

Table 6.2: The fifteen questions asked in the opening and closing surveys of the validation study.

arms (walking forward) and how to cause the robot to release (either by releasing pressure off the robot's back or by leaning back against the robot's arms). We chose this new standardized method of introducing users to the robot to more closely mimic how a user might learn how to hug the robot if he/she encountered it in the wild. We intentionally did not show the user performing any gestures on the robot in this instructional video to avoid biasing users toward this method of human-robot interaction. After watching this video, the user filled out an opening survey to document their initial impressions of the robot (Table 6.2). Importantly, this study included no formal practice hugs. Before the user began to interact with the robot physically, he/she watched the 30-second-long instructional video one more time as a refresher, since some pilot participants had forgotten what to do while completing the opening survey. The experimenter prompted them to "pay particular attention to the timing between the user and the robot."

Question
This robot behavior seemed (Unnatural – Natural)
These hug interactions felt (Awkward – Enjoyable)
These hugs made the robot seem (Socially Stupid – Socially Intelligent)
These hugs made the robot seem (Unfriendly – Friendly)

Table 6.3: The four questions asked after the introductory and concluding hug sessions (phases 1 and 3) in the validation study.

This study included three phases of physical interaction with HuggieBot 3.0; the robot always behaved autonomously, with no intervention by any experimenter and no changes to the program it was executing. The first phase was a natural hugging scenario. The users were given no additional instructions on how to hug the robot. They could perform as many hugs as they wanted, they could position their arms any way they wanted, and they could choose whether or not to perform any intra-hug gestures. They were simply told to hug the robot naturally several times, with a minimum of two hugs. We started the study with this open-ended phase to observe how users would naturally interact with the robot and to see whether they would discover the robot's ability to detect and respond to intra-hug gestures without experimenter prompting. Once they were satisfied with their introductory hugs, users were asked to fill out a single short survey for all the hugs they had performed, including answering the questions found in Table 6.3.

The second phase of the experiment included four hugs with intra-hug gestures conducted in a somewhat controlled manner. Before each hug, the user was instructed to perform a single gesture during the hug after the robot's arms had closed. They could perform the specified gesture as many times as they liked. The four gestures they were asked to perform were hold, rub, pat, and squeeze. We used an 4×4 balanced Latin square to counter-balance any effects of presentation order (Grant, 1948) and recruited participants in multiples of four to have an equal number of participants in all the presentation orders. After each hug, users were asked to fill out a short survey about the robot behavior they had just experienced in response to their gestures. This survey started with a free-response question asking for the user's "first impressions of this interaction." Then, they were asked to mark all the robot responses they experienced (options were robot arms staying still, robot arm moving vertically, robot arm tapping on back, and robot arms tightening hold). Finally, the participant used a sliding scale from 0 (hate) to 10 (love) to rate how they felt about the robot's response to the action they performed during the hug. Before moving on to the next part of the study, the user verbally explained their rating to the experimenter.

In the third phase of this study, users were once again asked to hug the robot naturally. They were allowed to hug the robot as many times as they liked. Afterward, users answered the same short questionnaire from the first phase, including the questions found in Table 6.3.

At the end of the third phase, users filled out a closing survey. The closing survey included the same questions from the opening survey (Table 6.2) plus questions asking users to rate the quality of the four different robot gestures they experienced during the experiment (hold, pat, rub, squeeze), as also done at the end of the action-response elicitation study. There were also two questions aimed at seeing how users responded to two new features we implemented: users rated the naturalness of the hug initiation method and the appropriateness of the robot's hand placement. Finally, users could provide additional free-form comments at the end.

6.8 VALIDATION STUDY – RESULTS

We analyzed the system's ability to perceive intra-hug gestures, the users' survey responses about the quality of their interactions, and user comments from all parts of the validation study to characterize HuggieBot 3.0's skill at autonomously and interactively hugging users. For reference, two annotated videos of participants performing active intra-hug gestures during their concluding hugs are included as supplementary material for this article; the annotations indicate the actions performed by the user, the actions perceived by the robot, and the gestures that the robot decides to execute in response. A video of a participant performing an extended hold during a hug in phase 2 is also included, annotated in the same way.

6.8.1 *Performance of the Perception Pipeline*

We analyzed data recorded by the robot during the second phase of the validation study to estimate the gesture perception pipeline's accuracy. The recorded data for each hug included timestamps, microphone voltage, and pressure values with a 45 Hz sampling rate, as well as the gestures detected every 10 samples. To determine accuracy, one of the authors visualized the

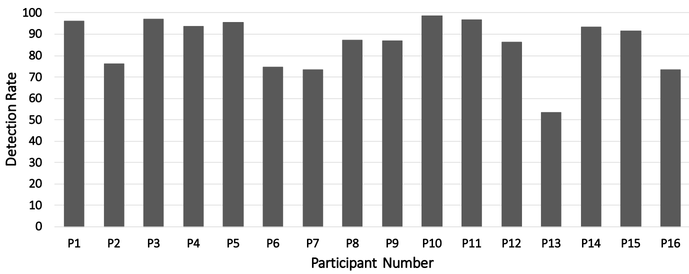


FIGURE 6.10: Gesture perception accuracy per participant.

data from the four hugs in the second phase of the experiment for all users (a total of 64 trials). As previously described, the participant was instructed to perform a specific gesture (e.g., squeeze) during each of these hugs. Since the users often interleaved their active intra-hug gestures with passive pauses, we calculated accuracy as the number of data points detected with the correct gesture or hold over the total number of detected gestures between the start and end of each hug. The mean detection accuracy for the 16 participants was 85.9% (standard deviation = 12.5%). This overall accuracy is comparable to the perception pipeline's accuracy (88%) on the data set collected in the action-response elicitation study. Figure 6.10 presents the perceptual accuracy for each participant. The gesture detection accuracy is above 86% for 11 participants, above 73% for four other participants (P2, P6, P7, P16), and at 53.5% for one participant (P13).

We examined the trials that had low detection rates and noted that sometimes the participants performed a gesture in an unexpected way. For example, for the squeeze gesture, P16 released the pressure on the chamber and applied it again instead of increasing the pressure (P16 – Squeeze, Figure 6.11). At other times, the participant applied little pressure (P13 – Squeeze, Figure 6.11) or squeezed the robot while performing another gesture (P6 – Rub, Figure 6.11). In some cases, the participant's accidental move or gesture, such as shifting their body against the front of the robot's torso, was detected as a rub (P13 – Hold, Figure 6.11). The algorithm also sometimes misclassified the start or end of a pat as rubbing.

6.8.2 *User Experience*

For all statistical analyses of the user responses to the questionnaires, we used an alpha value of $\alpha = 0.05$ to determine significance. To handle the

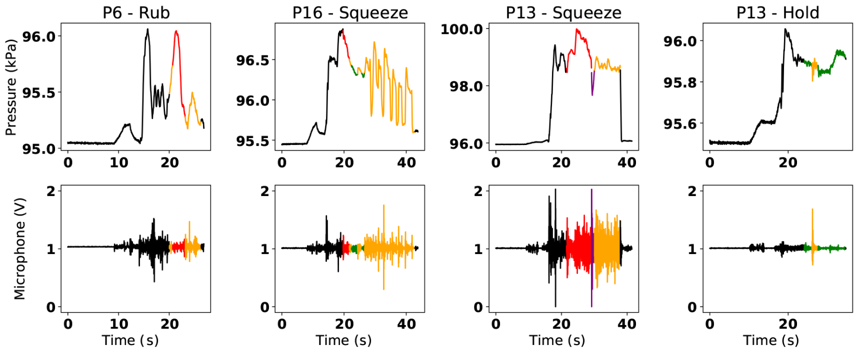


Figure 6.11: Example misclassified trials where the participants performed a gesture in an unexpected or uncommon way. Each panel's title gives the participant number and the action he/she was instructed to perform. The colored data points mark the time periods automatically classified as each of the four intra-hug gestures; following the color convention of Fig. 6.1, hold is green, rub is yellow, pat is purple, and squeeze is red.

problem of multiple comparisons and to lower the likelihood of a type I error, we used the Holm-Bonferroni method for alpha correction (Holm, 1978). This method adjusts the alpha value used to compare to the p-value to determine significance. First, the n obtained p-values are ordered from smallest to largest. Then, one compares the p_k value to the adjusted α value calculated with the following equation:

$$p_k < \frac{\alpha}{n+1-k} \tag{6.4}$$

where p_k is the probability that the tested comparison shows a significant difference, $\alpha = 0.05$ is the overall pre-specified alpha value, n is the number of comparisons being performed (the number of p-values being tested for significance), and k is the rank currently being evaluated. If p_k is smaller than α_k, the p-value is significant, and the null hypothesis is rejected; one increments k and moves on to the next p-value until p_k is not smaller than the adjusted value of α. When comparisons are significant, we report effect sizes using MATLAB 2019b's built-in function for Pearson's linear correlation coefficient: `ρ = corr(X)`, where X is the matrix of data being evaluated. The value of ρ signifies the strength of the bivariate relationship. $\rho = 0.1$ shows a low effect size, $\rho = 0.3$ indicates a medium effect size, and $\rho = 0.5$ or above signals a high effect size.

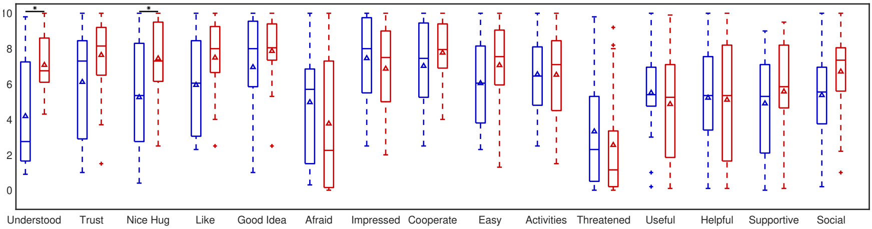


FIGURE 6.12: Box plots comparing the responses to the opening (blue) and closing (red) survey questions given in Table 6.2. The top and bottom of each box represent the 25th and 75th percentile responses, respectively, while the line in the center marks the median, and the triangle shows the mean. The lines extending past the boxes show the farthest data points not considered outliers. The + marks indicate outliers. The two black lines with stars at the top of the graph indicate statistically significant differences.

Box plots of the user responses to the opening and closing survey questions from Table 6.2 are shown in Fig. 6.12. In this study, answers were submitted on a continuous sliding scale from 0 (disagree) to 10 (agree), so a paired t-test comparison of the opening and closing survey was conducted after verifying the data were normally distributed. After applying the Holm-Bonferroni method for alpha correction (with a $n = 15$ for the fifteen questions in the survey), we found that users felt significantly more understood by the robot ($p = 0.0023$, $\rho = 0.21$) and felt the robot was significantly nicer to hug ($p = 0.0035$, $\rho = 0.59$) after the experiment. Three other comparisons approached significance: users liking the presence of the robot ($p = 0.0084$, $\rho = 0.72$), thinking the robot could support them ($p = 0.0372$, $\rho = 0.92$), and viewing the robot as a social agent ($p = 0.0208$, $\rho = 0.68$).

The first and third phases of the study involved asking the participants to perform several natural hugs with the robot. The responses to the four questions from Table 6.3 asked after the introductory hugs (phase 1) and the concluding hugs (phase 3) can be seen in Fig. 6.13. These responses were also analyzed using a paired t-test comparison and a Holm-Bonferroni alpha adjustment. We found that users did not initially find the robot's hugging behavior very natural, but their opinion of it significantly improved by the end of the study ($p = 0.0138$, $\rho = 0.47$). At the end, users also found the robot hugs significantly more enjoyable ($p = 0.0028$, $\rho = 0.77$). Finally,

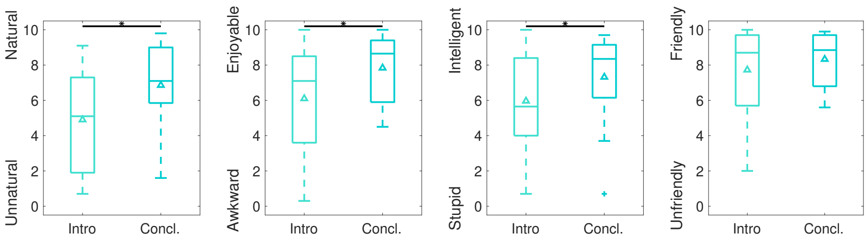


FIGURE 6.13: A box plot comparison of the responses to the four survey questions about the introductory hugs (light turquoise) and concluding hugs (dark turquoise), as listed in Table 6.3.

after hugging the robot repeatedly, users found the robot significantly more socially intelligent ($p = 0.0049$, $\rho = 0.84$) than their initial impressions. The users did not significantly change their opinion about the robot's friendliness over the course of the study; it was already rather highly rated after the introductory hugs.

We also investigated the average hug duration and the average number of gestures users performed during the introductory and concluding hug phases (Fig. 6.14). For the introductory hug phase, the average hug duration was 22.7±12.6 seconds, where 12.6 seconds is the standard deviation. The concluding hug phase had an average hug duration of 25.3±11.0 seconds. The average number of active user gestures (rub, pat, squeeze) detected during the introductory hug phase was 1.41±1.84. The average number of active gestures detected during the concluding hug phase was 4.03±3.95. We ran a paired t-test on the average hug duration and the average number of gestures in these two phases. Although twelve of the sixteen users engaged in longer hugs with the robot during the concluding phase than during the introductory phase, we did not find a significant difference for hug duration. However, we did find a significant difference for the number of gestures performed ($p < 0.001$, $\rho = 0.89$), with significantly more gestures performed during the concluding hugs.

Users experienced four hugs during phase two of the study, performing a specific gesture as many times as they liked in each hug. The robot used its perceptual pipeline and probabilistic behavior algorithm to autonomously decide how to respond to each of the user's actions. Figure 6.15 shows a box plot of the user ratings for the overall robot responses for each performed user action. For all user actions, the average rating of the robot's combined response is around eight out of ten with a standard deviation of

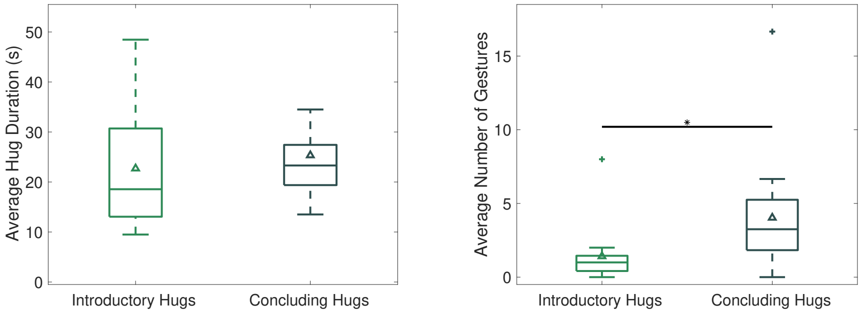


FIGURE 6.14: A box plot comparison of the average duration (left subplot) and the average number of intra-hug gestures (right subplot) of the hugs during the introductory hug phase (pale green) and the concluding hug phase (dark green) of the validation study. The black line with a stars indicates a statistically significant pairwise difference.

around 1.5 ($r_{\text{hold}} = 7.90 \pm 1.77$, $r_{\text{rub}} = 7.94 \pm 1.64$, $r_{\text{pat}} = 7.82 \pm 1.35$, and $r_{\text{squeeze}} = 7.96 \pm 1.42$), indicating that the robot's responses were perceived very positively by users.

At the end of the experiment, users were asked to evaluate the quality with which the robot performed the four different intra-hug gestures. This question was presented and framed the same way as in the action-response study, for continuity and comparison. A matrix showing the average rating on a color scale and each user's individual rating in dots can be seen in Fig. 6.16. Figure 6.17 shows side-by-side comparisons of the gesture quality ratings obtained in the original action-response elicitation study and those from the validation study. We then compared the quality ratings of the four different robot actions between the two studies using unpaired t-tests. We had a different number of participants and an entirely separate population. A Holm-Bonferroni alpha adjustment with $n = 4$ did not show any significant differences in the perceived quality of the robot gestures.

Fig. 6.18 reports how the users rated the naturalness of the hug initiation process and the appropriateness of the robot's hand placement on their back. The average rating of the hug initiation 6.94±2.22 on a scale from unnatural (0) to natural (10). The average rating of the robot's hand placement was 8.40±2.02 on a scale from inappropriate (0) to appropriate (10). One outlier rated robot hand placement a 4 out of 10. This user was one of the tallest users we tested, with a height of 1.83 m. Two other users had the same

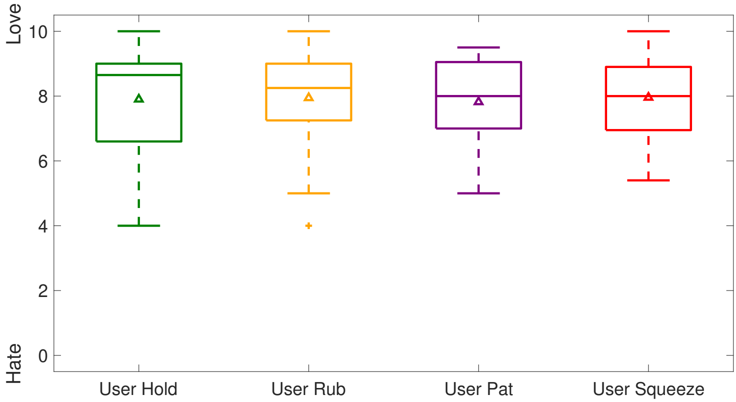


FIGURE 6.15: A comparison of the user ratings of the robot's autonomous responses to the four different intra-hug actions performed by users in the second phase of the validation study.

height as this user and rated it 6.3 and 9 out of 10. This user tended not to stand straight when the robot was estimating his height, and thus, HuggieBot 3.0 thought he was shorter than he was, which resulted in the robot's arms occasionally being placed lower than the user would have preferred.

6.8.3 *User Comments*

Once again, some of the most informative results come straight from the users in the form of free-response written and verbal comments. Half of the users commented explicitly on the naturalness and ease of initiating the hug with the robot. 31.25% of the 16 users specifically commented that the "timing was well synchronized/done perfectly on time" (P10, P13) between initiating a hug and the arms closing. One user (6.25%) initially felt the arms closed too slowly. However, by the end of the experiment, this user stated the timing was "natural and comfortable" (P12, P15) for them. Interestingly, two users (12.5%) thought the robot's arms closed slightly faster than with people and felt they had to walk slightly faster than usual. Regarding the robot's hand placement, 75% of users mentioned that the location of the robot's hand was "good" and "well-placed exactly where [they] want it" (P15) at all times. Users specifically mentioned that at no

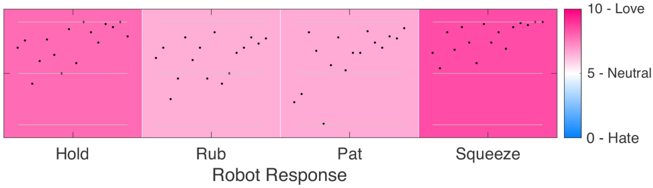


Figure 6.16: A matrix showing the user ratings of the quality of the various robot responses from the validation study, following the visualization approach used in fig. 6.3.

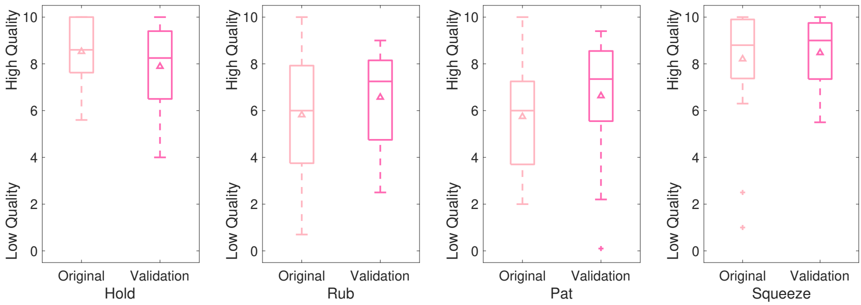


Figure 6.17: A comparison of the responses to the quality of the robot's gestures during the original action-response elicitation study (pale pink) and the validation study (dark pink).

point did it make them feel uncomfortable, which is an improvement from the user comments from the action-response elicitation study.

After the introductory hugs, six users (37.5% of 16) mentioned that hugging the robot felt "strange" (P2) or "weird" (P6, P7, P9) because it was their first time interacting with a robot, and they were not sure what to expect or how the robot would respond. While these users initially found hugging a robot strange, that does not mean they did not enjoy their experience. 50% of all the users mentioned enjoying their first group of robot hugs, particularly commenting on the robot's "warmth" (P4, P5, P10, P12, P16), "friendliness" (P2, P7, P12), and "comfortable appearance" (P2, P4, P6, P7). Some users even went as far as to say "I enjoyed the experience as I am away from my family and a warm hug is always comforting" (P4). Only three users (18.75%) mentioned feeling uncomfortable or nervous initially.

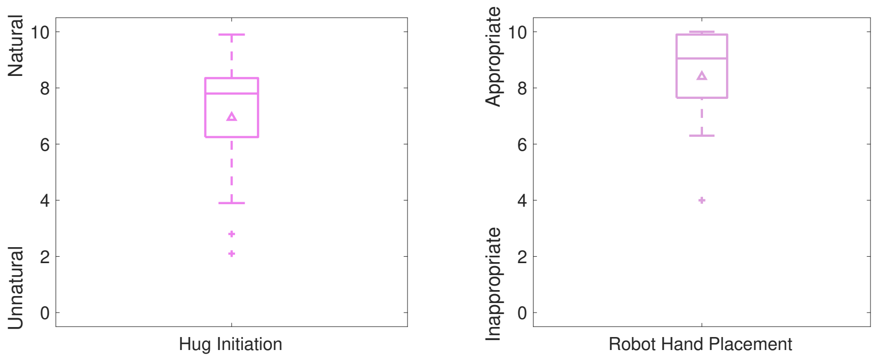


Figure 6.18: Box plots showing user ratings of the naturalness of the hug initiation process (left) and the appropriateness of the robot hand placement (right).

After the concluding hugs session, user comments were much more positive. Fifteen out of our sixteen participants (93.75%) shared positive comments mentioning "great overall experience" (P3, P4) and that they "liked the robot's responses" (P4, P5, P6, P12, P15, P16). Users mentioned that they found the hugs to be both "comfortable" (P2, P4, P6, P7) and "natural" (P4, P7 P10). The one user who did not have as positive comments as the rest mentioned that "the experience did not feel natural, but it was fun to test how a robot can interact with a human in this way" (P8). Three users who were initially tentative about the robot in their comments about the introductory hugs stated how comfortable they eventually found the robot (18.75%). Six users (37.5%) were "amazed" (P3, P7) by the experience and came to think of the robot as a "friend" (P2, P7, P14) by the end of the concluding hugs phase.

We also asked users to share their opinions about the overall experience at the end of the entire study. Many users (62.5%, 10 out of 16) shared that they felt the entire experiment was a "great experience" or that they "loved hugging the robot" (P5, P6). Two users (12.5%) mentioned that while they enjoyed the interaction, they did not understand the purpose of such a robot as they felt that human hugs are "irreplaceable" (P12, P13). Several of our users (37.5%, 6 of 16) mentioned how valuable they found this robot especially given the current COVID-19 pandemic, mentioning that "emotional and mental health are also important" (P7) but are often forgotten or not addressed.

6.9 DISCUSSION AND LIMITATIONS

6.9.1 *Discussion*

This chapter investigates the previously unstudied phenomenon of intra-hug gestures during hugs between a human user and an autonomous robot. We found users to be positively interested in hugging a robot that can both respond to their gestures and proactively perform gestures on its own. The results support all six new of the tenets we proposed for hugging robots.

T4: HUG INITIATION First, our revised version of the fourth tenet states that a hugging robot should autonomously initiate a hug when it detects a user in its personal space by inviting the user for a hug; it should then wait for the user to begin walking toward it before closing its arms to ensure a consensual and synchronous hugging experience, as done by HuggieBot 3.0. The previous evaluation of HuggieBot 2.0 did not find a statistically significant user preference between a hug initiated with a button press or a hug initiated via computer vision detection of the user's approach (Block, Christen, et al., 2021). Users gave our new two-phase hug initiation method an average naturalness rating of almost a 7 out of 10. All users were able to initiate hugs with the robot after watching a simple video demonstration twice, without requiring the detailed instructions that were previously provided. Thus, we conclude that the new hug initiation method is an improvement over the previously tested methods. We also had pilot-tested an alternative three-step hug initiation method that was harder for users to master and received an average naturalness rating of only 4.62 out of 10 from 12 pilot participants. Our observations repeatedly indicate that consensual and synchronous hug initiation is indeed important. Although it still has room for improvement, rather than forcing the user to hurry up and wait, the new two-phase method more closely mimics how hugs occur between humans and lets the user decide when to start hugging the robot.

T7: ADAPTATION TO USER HEIGHT The next new tenet states that a pleasant hugging robot should also perceive the user's height and adapt its arm positions to contact the user at appropriate locations. We thus extended our platform's perceptual capabilities beyond detecting a user's approach so that HuggieBot 3.0 attempts to embrace the user at the proper height – not too high, and not too low. Our newly developed height adjustment system was created in direct response to user feedback about the inappropriateness of HuggieBot 2.0's hand placement in the action-response elicitation study. The average rating for the appropriateness of HuggieBot3.0's hand placement's was an 8.4 out of 10, with several users giving it the highest rating possible; these ratings indicate that the proposed approach usually succeeds at adjusting for user height. We also believe that the non-significant increases in the quality ratings for the rub and pat gestures (which we did not adjust other than the location of the robot's left hand) can probably be attributed to this improved placement. Thus, we conclude that users prefer hugging a robot that adjusts its arm placement to match their height.

T8: GESTURE PERCEPTION The next tenet centers on enabling a hugging robot to accurately and reliably detect and classify user gestures applied to its torso in real time, regardless of the user's hand placement. Both of our user studies demonstrated the excellent haptic sensing capabilities of the pressure sensor and microphone inside the inflated chamber of HuggieChest; simple signal-processing and machine-learning techniques were able to detect and classify contacts very well, even when the pipeline was transferred to new sensing hardware and adapted with limited new training data. As our users used various hand positions on the robot (both arms below the robot's arms, both above, or one above and one below), we found that this non-localized haptic sensing system works well regardless of user hand placement on the back chamber. Based on the average gesture detection accuracy of 86% for the 16 participants in the validation study, along with the positive opinions users shared when the robot responded to their gestures in both studies, we believe the results support the validity of this tenet. In further exploring cases when the detection algorithm did not perform well, we found that users frequently performed the gesture in an unexpected or uncommon way; in some cases, these variations may have come from the user's limited vocabulary for intra-hug gestures in English. As a surprising benefit, however, the perception pipeline was able to detect some rubs and pats performed at the same time as a squeeze, even though it had not been trained to do so.

T9: FAST RESPONSE The ninth tenet simply states that users like a robot that responds quickly to their intra-hug gestures. As seen in Fig. 6.3, when a user performed an intra-hug gesture on the robot's back, and the robot did not respond, users perceived this as a neutral robot behavior on average. In their written and verbal comments to the experimenter, users indicated they did not feel like the robot "understood" them, "knew [they were] there," or "wanted to support/comfort [them]." Users clearly preferred when the robot indicated that it knew the user had performed an intra-hug gesture and responded quickly in some way. When we were piloting the validation study, we noticed that a technical error occasionally caused the sampling rate of the microphone and pressure sensor to drop to about half the normal value. We found that when users interacted with the robot in this condition, the delay was highly noticeable and detracted from the user experience. Several pilot participants mentioned that "it felt like the robot was performing random actions at random intervals, not in response to anything I was doing." During our final validation study, when the sampling rate issue was fixed, the robot responded quickly to user actions, and our participants were delighted. They repeatedly performed the same gesture to experience the robot response again. Users would often comment to the experimenter during the hug itself, saying "it tapped me back!" (P4), or remarking after the hug that "every time I did an action, it noticed and did something back to me!" (P16) because they were so pleasantly surprised at the responsiveness of the robot.

T10: RESPONSE VARIETY The next tenet states that hugging robots should adopt a gesture response paradigm that blends user preferences with slight variety and spontaneity. When starting this project, we believed that hugging robots should always reciprocate the same intra-hug gesture the user had performed. The results from the action-response elicitation study surprised us by showing that rote reciprocation is not expected and would not be perceived in a fully positive way. If users preferred gesture reciprocation, we would see a dark pink diagonal in 6.3. Instead, we see a slight preference for a robot to respond to any user action with a squeeze. Speaking with our users showed us that they appreciate variety in robot responses. Something about the unpredictability of the response leads users to feel it is more "alive." Users also mentioned that having the robot respond with the same action as the user performed feels "too mechanical," because based on the input you know exactly what output you will receive. The results from the action-response elicitation study thus support this tenet, as do the very positive user reactions to the resulting robot behavior algorithm tested in the validation study. We believe the slightly spontaneous robot hugging behavior enabled by our simple probabilistic behavior paradigm (equation (6.3)) succeeds at blending user preferences with spontaneity to reasonably match natural human exchanges of intra-hug gestures. The behavior algorithm's tendency to prefer exploration versus exploitation can also easily be adjusted by changing the value of the exponent m.

T11: PROACTIVE ROBOT GESTURES This tenet states that hugging robots should occasionally provide unprompted proactive affective social touch to the user through intra-hug gestures. The findings of T. L. Chen et al. (2014) made us initially hypothesize that users would dislike robot-initiated affective social touch delivered via unprompted intra-hug gestures; their users reacted negatively when a robot attempted to comfort them by touch but did not mind functional contact from the robot. The findings of our two studies explicitly contradict this hypothesis and support tenet 11. We were so surprised by these ratings during the action-response elicitation study that after the user had finished explaining their ratings to the experimenter, she asked the follow-up question, "so just to clarify, it did not bother you that you did nothing and the robot unprompted started rubbing/patting/squeezing you?" Users confirmed that not only did they not mind this robot behavior, but they also *enjoyed and appreciated it*. Users indicated that while in the other cases, the robot would respond to their gestures, here, they felt the robot was comforting them. In these cases, many users commented that they felt the robot's emotions and feelings and that it cared more about them when it chose on its own to perform a gesture, rather than just responding. Although more work needs to be done to confirm this positive finding, it seems that appropriately framed robot-initiated affective touch may be key to creating robots that can provide good emotional support to human users.

How can we grapple with the seemingly conflicting findings between our work and T. L. Chen et al. (2014)? We believe these results are not as different as they may appear. The users in our studies *agreed to enter into a hug with a robot*, so we believe they also felt at least partially responsible for initiating the affective touch that occurred during the resulting hug. Once this initial boundary is broken, we believe users are more receptive to proactive robot affective touch, for example, a rub, pat, or squeeze. Users in all of our studies have appreciated that HuggieBot *politely asked* them for a hug, thereby allowing them to agree to this affective touch. Many users even responded affirmatively to the robot every time it asked the question, even though they knew it never listened to their answer. By changing the hug initiation method to be prompted by the robot lifting its arms for a hug and asking "Can I have a hug, please" and then waiting with its arms outstretched for the user to approach, we further put the initiation of the affective social touch on the user, solidifying that it is their choice to enter the hug. We believe user initiation is key to acceptance of future social, affective touch from a robot. We therefore firmly believe tenet 11's statement that robots can evoke user feelings that the robot is alive and caring by occasionally providing unprompted affective touch to the user, as delivered by HuggieBot 3.0 through intra-hug gestures.

6.9.2 *Limitations*

While the research described in this chapter presents several key contributions to robotic hugging and broader social-physical human-robot interaction, we nevertheless acknowledge several limitations of our work.

The first limitation is the somewhat artificial methodology of our studies. We recognize the importance of conducting in-the-wild studies for human-robot interaction research. By conducting these studies in a laboratory environment, we have a self-selection bias of our participant pool. Only users who were interested in hugging a robot chose to participate in the reported studies. Unfortunately, due to the current COVID-19 pandemic, lab studies were the safest way to conduct research on hugging robots. We were able to screen participants for potential health risks and thoroughly sanitize the robot between participants. For the validation study, we changed the robot's introduction from verbal instructions from the experimenter to having the user watch a simple video of another user hugging the robot. This video introduction was meant to mimic how users would learn to use the robot in the wild.

We have also identified two main limitations of the Kinova JACO arms used in both HuggieBot 2.0 and 3.0. While we adjusted the hug initiation process to accommodate this limitation better, the Kinova JACO arms just cannot move fast enough to mimic the speed of a human's arms closing. These arms were selected for safety reasons, and this speed limit was considered during that choice. After extensive user testing, we have found that the speed limit causes a real limitation on the naturalness of the user experience because users have to wait several seconds before the arms have fully closed around them. A second limitation of these arms is that repeated small movements of the first, second and third joints (shoulder lift, shoulder pan, and elbow flex, respectively) occasionally result in a sudden short but fast jerking movement, which startles the user. This phenomenon occurs only rarely during repeated rubs or pats. When this issue occurred during the two reported studies, we immediately commanded the robot to release the user, checked that he/she was okay, verified whether he/she wanted to continue, discounted the trial with the malfunction, and restarted the trial. Because the sudden motion is very small, this technical glitch never hurt a user. However, it is likely that it negatively affected some user ratings of HuggieBot as a whole, as well as the robot's ability to perform rubs and pats.

Another limitation is that both of our studies asked the user to perform intra-hug gestures somewhat artificially. After placing their hands on the robot's back, users had to wait for the robot's arms to close fully before performing a gesture. This pause was used to collect baseline measurements for the microphone and pressure signals so that the real-time perception pipeline could determine what gestures the user subsequently performed. We found that many users naturally wanted to start performing the gesture immediately after beginning the hug, regardless of the robot's arm movements. To collect data and then test our algorithm's accuracy, we also asked users to perform only one gesture per hug, though they could perform the gesture repeatedly if they chose. This restriction was also somewhat unnatural. We found that many users naturally wanted to combine gestures. We added the natural hug scenarios in phase 1 and phase 3 of the validation study to address this limitation.

Our action-response elicitation study challenged users with the difficult task of separating the appropriateness of the robot's response from the quality with which HuggieBot 2.0 performed the gesture. We had them explain their ratings to the experimenter to ensure they understood the distinction and were answering the question correctly. Nonetheless, the robot's gesture quality probably affected other user ratings. Users experienced a similar challenge in the validation study, where we again asked them to separately rate the robot responses from the quality of the robot's gestures; gesture quality also thus probably affected these results. Another limitation from the action-response elicitation study is that we did not ask users to rate the naturalness of the hug initiation process or the appropriateness of the robot's hand placement. We had not realized these aspects of HuggieBot 2.0's behavior would garner negative comments and thus need to be adjusted for the new version of the platform. Thus, we had to rely on written and verbal comments to evaluate the effects of these changes.

A final limitation we observed is that the training data used to train our detection and classification algorithm came from a combination of data from two different physical torsos. Had we only used the original torso with data from 16 hugs from 32 different people, our detection accuracy could be higher. Unfortunately, we had not anticipated the strength with which some users would squeeze our robot, which ultimately created severe leaks that were not repairable. Once we realized we would have to make a new torso, we decided it would be best to make the chambers a uniform size and shape. This change in the torso's physical construction meant that the data we collected was not able to transfer on its own without some additional data collection. While we still used the original data, we had to supplement it with several new trials on the new torso. Due to difficulty recruiting during COVID-19 restrictions, we decided to augment the previous dataset with new trials from a few participants instead of collecting a whole new dataset. Finally, while using a larger window size would increase our detection method's accuracy, it also slowed down the robot's response in real time. The delay was quite apparent to our users, and thus we opted for a smaller window with lower accuracy to maintain realistic real-time responses.

6.10 CONCLUSION

This chapter started by collecting a large data set that shows the characteristic microphone and pressure signals for 32 diverse users performing four intra-hug gestures (hold, rub, pat, and squeeze) on the inflated torso of HuggieBot 2.0. We used these recordings to create a perception pipeline that detects and identifies these different gestures in real time. Ratings and comments in reaction to how the robot responded showed that users do not want a robot that mimics their gestures back to them; instead, they want a robot that responds quickly and naturally to their gestures with some level of unpredictability similar to the choices made by a human hugging partner. We thus developed a behavior algorithm that uses conditional probabilities based on user ratings to determine how our robot should respond after detecting a particular user action during a hug, distinguishing between discrete active gestures, modal active gestures, and passive gestures. We also made several critical changes to our robot platform, including changing the method of hug initiation, adjusting the robot's arm positions to the estimated height of the user, constructing a new robot torso, and improving the quality of the robot's embrace. We tested this upgraded version (HuggieBot 3.0) together with its new perceiving and acting skills on a new set of 16 diverse users who had not previously interacted with this robot. Users were generally very pleased with the robot's responses to their actions. The platform changes seemed to improve the quality of the interaction, and the perception and behavior approaches that we developed worked very well.

Therefore, we conclude that hugging robots should be able to perceive and respond quickly to intra-hug gestures from users.

7

EMOTIONAL AND PHYSIOLOGICAL EFFECTS OF ROBOT HUGS

A hug a day keeps the demons at bay.
— German Proverb

7.1 INTRODUCTION

Oxytocin and cortisol are two key neurohormones produced within the human body; oxytocin also serves as a paracrine hormone with receptors spread throughout the body (Kubzansky et al., 2012). While popular media may refer to oxytocin as the "love hormone," oxytocin is associated with positive emotion, social bonding, and feelings of well-being (anxiolytic). Interestingly, oxytocin is also associated with increased pain tolerance, thus allowing us to endure more stress or pain than we otherwise might be able to handle (Uvnäs-Moberg, Handlin, & Petersson, 2014). Cortisol is commonly referred to as the "stress hormone" because it triggers the body's "fight or flight" response. When we feel a threat, cortisol heightens the nervous system and motor reactions needed to survive (pupils dilate, salivation decreases, airways relax/expand, heart rate increases, etc.) At the same time, cortisol overrides any calmer functions the body determines to be non-essential for survival.

While much research has been conducted to characterize how affective, social interaction affects these two hormones (Uvnäs-Moberg, Handlin, & Petersson, 2014), little has been researched to see whether and how they are affected by social touch provided by a robotic agent. Not only is social touch helpful both physically and emotionally, but the absence of social touch can have detrimental effects (Morrison & Gore, 2010; Neira & Barber, 2014). The current global COVID-19 health crisis and the resulting need for physical distancing have shown that people in isolation miss hugs and emotional connections with their loved ones. We believe a hugging robot may help people bridge this gap of staying physically apart while still having their emotional needs met through a safe form of social touch. In this chapter, we report the first study to measure the extent to which a hug from a robot

FIGURE 7.1: These four images show the waiting-for-a-hug pose and the invitation-for-a-hug pose for the human hug condition (left) and the robotic human hug condition (right).

can affect a user's emotional and physiological response after stress and directly compare it to the effects of a human hug.

7.2 CONDITIONS AND HYPOTHESES

Until this point, most of our evidence supporting the benefits of robot hugs has been anecdotal comments or self-reported survey responses. This chapter seeks to gain alternative perspectives to understand the extent to which active robot hugs can benefit users by adding physiological measures of heart rate, spitting time, salivary oxytocin, and salivary cortisol. We define a robot hug to be one where the robotic hugging agent actively hugs the user back by wrapping its arms around the user; in contrast, a passive robot hug is one in which the robot only receives a hug, as in Sumioka et al. (2013), and it does not move in response.

We compared five different hugging conditions, varying both the hugging agent and the hugging agent's behavior: (A) a human hug, (B) a robotic human hug, (C) a robot hug, (D) a passive robot hug, and (E) no hug.

(A) **Human Hug:** This condition took place in the Human Room.

Appearance: The human agent's appearance can be seen in Figure 7.1. The woman pictured is a professional actress who was hired to play the role of the human hugging agent throughout this study. When wearing shoes, she is the same height as HuggieBot. In this condition, the human hugging agent wore the purple robe and grey sweatshirt. She also wore padding around her stomach to reduce her waist-to-hip ratio, so she had similar body proportions as the robot. She wore an FFP2 mask for COVID-19 safety during this close interaction. Her hands and eyes were visible. In this condition, she made eye contact with the user when inviting him/her for a hug, and after the hug.

Movement: She moved at a normal speed. When she lifted her arms to invite the user for a hug, they were between waist height and shoulder height. The waiting pose and invitation pose can be seen in Fig. 7.1. She responded to any intra-hug gestures performed by the user naturally with a squeeze. When the user released her, she removed her arms at a normal speed and dropped them to her sides.

(B) **Robotic Human Hug:** This condition took place in the Human Room.

Appearance: In this condition, the human hugging agent wore the same purple robe, grey sweatshirt, stomach padding, and FFP2 mask as described in the above condition. The robotic human agent's appearance can be seen in Figure 7.1. In addition, on top of the FFP2 mask, she wore a clear face shield in order to mimic the screen of the robot's face. In this condition, she wore the same grey mittens as the robot to cover her hands. During this interaction, the human hugging agent did not make eye contact with the user. She looked at a point near the user's head without looking into the user's eyes. Thus, in this condition, we obscure both the human hugging agent's eyes and hands.

Movement: The human hugging agent matched her movement to that of the robotic hugging agent. She moved slightly slower than she did in the human hug condition to match the speed of the robot's arms. When she lifted her arms to invite the user for a hug, she lifted them higher than she naturally would, and she raised her right arm higher than her left, again to match the robot's pose. The robotic human agent's poses can be seen in Fig. 7.1. She closed her arms in a slow and controlled manner, so the left arm finished closing before the right arm, similarly to how the robot behaves. If the user did not perform any intra-hug gestures, she pro-actively performed a two-second squeeze after 1.5 seconds of holding the hug. If the user performed either a rub or a pat on her back, she responded with a two-second squeeze. Finally, if the user squeezed her, she reciprocated by holding the squeeze until the user released. When the user released the hug, she released it more slowly than natural for a person, using a speed that matched the robot. When she finished moving, she held

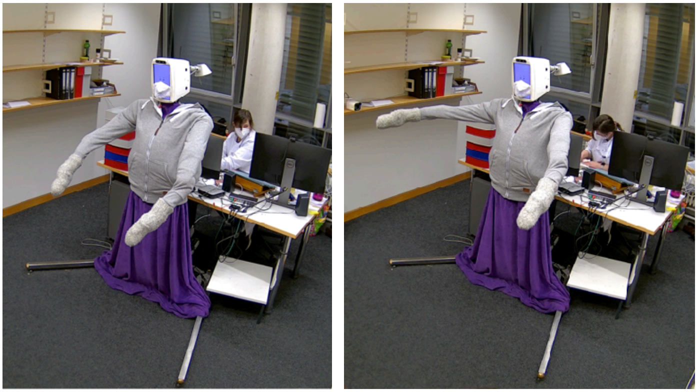


FIGURE 7.2: HuggieBot 4.0, wearing an FFP2 mask, in the "waiting" and "invitation" poses.

her arms in a ballet first-position to match the robot when waiting for a user.

(C) **Robot Hug:** This condition took place in the Robot Room, which is identical in size to the Human Room and is located only a few doors away on the same hallway.

Appearance: The robot wears the inflated HuggieChest torso, covered by two heating pads (one on the chest and one on the back), a robe, and a sweatshirt. The purple robe and grey sweatshirt are duplicates of the ones the human hugging agent wears in the two conditions described above. About ten copies of both garments were purchased to allow washing after each participant for both the human and the robot. The robotic agent's appearance can be seen in Figure 7.2. Before the user entered the room, we turned on both heating pads, so the robot was both warm and soft for the interaction. The robot has a purple face and cycles through four different smiling and blinking faces to give the appearance it is "alive," based on Fitter and Kuchenbecker (2016a). Finally, the robot's end effectors are covered with grey mittens that match those worn by the human in the robotic human hug condition described above.

Movement: The robot moved to the best of its abilities in this condition. Upon detecting a user waiting for a hug, it lifted its arms and told the user it was ready for a hug. The robot waited with its arms outstretched until the user began walking forward, which triggered the robot to start closing its arms. The robotic agent's poses can be seen in Fig. 7.2. If the user stopped moving forward or backed up, the robot would reopen its arms to avoid closing on air (without embracing the user). As described in Chapter 6, the robot also used the camera to estimate the user's height and lift or lower its arms to contact the user in the middle of the back (not too high or too low). The robot used its novel adaptive hugging as a form of grasping to create a secure, customized embrace for each user. The robot used the HuggieChest and detection and classification algorithm to determine if the user made any intra-hug gestures. It also used its probabilistic behavioral algorithm to respond to these gestures with squeezes (timed if the user rubbed or patted, and of reciprocal duration if the user squeezed). If the robot did not detect any actions for 1.5 seconds, the robot pro-actively performed a timed squeeze. When the user either released pressure from the robot's back or pushed against its arms, thus indicating their desire to be released, the robot opened its arms and let the user go.

(D) **Passive Robot Hug:** This condition took place in the Robot Room.

Appearance: In this condition, the robot's appearance is exactly the same as it was in the Robot Hug condition.The robotic agent's appearance can be seen in Figure 7.2. The difference between these two conditions is in the robot behavior/movement.

Movement: When it detected a user present, the robot lifted its arms and told the user it was ready for a hug, just like in the Robot Hug condition. The robotic agent's poses can be seen in Fig. 7.2. The robot still used its visual perception to initiate the arms closing, but the height of the robot's arms did not change according to the user's height. While the robot's arms closed, they stopped in the air to avoid making contact with the user's body. The arms closed $15°$ from the invitation for a hug pose. This embrace was large enough that the arms did not touch even our largest user. In this condition, the robot also did not respond to intra-hug gestures performed by the user. The robot did respond and release when the user either removed their hands from the robot's back or pushed against the robot's arms.

(E) **No Hug:** This condition took place in the Human Room. The appearance of the empty room can be seen in Figure 7.3. In this condition,

Figure 7.3: The empty room users experienced in the no hug condition.

there was no hugging agent in the room with the user and the experimenter. Otherwise, the room was set up the same.

These conditions were chosen to allow us to test the following four hypotheses:

H1. Any hug is better than no hug at all. Therefore, any of the four conditions with a hug will provide greater enjoyment (self-reported survey responses) and benefits (increased oxytocin, decreased cortisol, and faster heart rate reduction) than no hug.

H2. A human hug is irreplaceable; this condition will provide greater enjoyment and physiological benefits than all other tested interventions, including an active robot hug.

H3. Being actively hugged by a partner increases the enjoyment and physiological benefits a human reaps from a hug. Thus, an active robot hug will provide greater enjoyment and benefits than a passive robot hug.

H4. As the robot hug and the robotic human hug have the same movement and appearance, with the only difference being the agent itself, users will prefer the robot hug over the robotic human hug because of the lack of social judgment.

To further explore the effects of active robot hugs, we had all study participants exchange a hug with HuggieBot in the robotic hug condition about 35 minutes after a stressful experience. We expect users to have a variety of stress levels at this point in time, depending on both their individual characteristics and the condition to which they were assigned for the first intervention. We added this second hugging intervention to test the following hypothesis:

H5. Experiencing an active robot hug about 35 minutes after a stressful experience will positively affect all participants.

7.3 SYSTEM

As mentioned earlier, in this study, we had two hugging agents who each performed in two different ways. In this section, we discuss the robotic hugging agent and then the human hugging agent. Then, we describe the different behaviors of the hugging agents in their different conditions. Finally, we explicitly state the upgrades we implemented to the robot from Chapter 6.

7.3.1 *Robotic Hugging Agent*

This experiment builds upon the validated custom hugging robot HuggieBot 3.0 (Block, Seifi, et al., 2021) (Fig. 7.2), which is closely based on HuggieBot 2.0 (Block, Christen, et al., 2021). The robot is 1.75 m tall with proportions roughly similar to a human. The robot's custom torso, HuggieChest, both softens the robot and senses the user's contacts on the robot's back via dual sensors: an Adafruit BME680 barometric pressure sensor and an Adafruit electret microphone amplifier MAX4466, which are connected to an Arduino Mega. The micro-controller sends the data to ROS (Robot Operating System) over serial at approximately 45 Hz. On top of each of the torso's chambers is a Thermophore MaxHeat Deep-Heat Therapy heating pad (35.6 cm $\times$ 68.6 cm) to warm the robot to a temperature similar to a human body. The robot's arms are two six-degree-of-freedom Kinova JACO arms mounted horizontally on a custom metal frame. Though they are slightly longer than human arms, they are anthropomorphic and quiet. Foam pads cover the arms for user comfort and safety. A Dell OptiPlex 7050 minicomputer controls the entire robot from inside the robot's head. Also in the head is the robot's face screen, a speaker, and an Intel RealSense depth-sensing camera used to detect an approaching user. The robot wears a purple robe with a gray sweatshirt and gray mittens, and it has a female voice in this study.

7.3.2 *Human Hugging Agent*

To eliminate variables, we employed a female actor of similar height and build as HuggieBot. As described in Section 7.2, she was dressed the same as the robot and stood in the same relative location in a neighboring room closely matched to HuggieBot's room.

We chose to hire a female actor and to gender our robot as female based on several lines of research. First, touch is the primary mechanism for providing life sustaining care for an infant. Babies need physical touch and skin-to-skin contact to eat, gain weight, and thrive (Feldman, Rosenthal, & Eidelman, 2014). Infants respond to touch from caregivers with oxytocin release (Insel, 1992; Keverne & Kendrick, 1992; Meaney, 2001) setting up their pattern of physiologic organization, cognitive control, and responses including stressors across life (Feldman, 2017, 2020; Yaniv et al., 2021).

Around the world, women are typically the primary caregivers, and 70% of people develop secure attachments to their primary caregiver (Ainsworth, 1982). The 30% who do not develop secure attachments do so, in part because they do not receive reliable, contingent care when needed, particularly during early developmental sensitive periods. Over the first two to three years of formative social, emotional, cognitive, behavioral, and personal development, if their caregiver(s) do not respond to their needs and crying with responsive care communicated via comforting touches, their Hypothalamic Pituitary Adrenal (HPA) Axis their stress response and emotional balance are negatively affected (Juruena et al., 2020). Oxytocin's dynamic role in counter-balancing cortisol's physiological arousal includes both oxytocin contingent release as well as patent oxytocin receptors to process oxytocin, hypothesized to be nature's medicine (Carter et al., 2020) Generally, therefore, the majority of people across cultures have secure attachments to an older female caregiver in their life (Santrock, Deater-Deckard, & Lansford, 2021). We wanted to draw on this connection, comfort, and the positive associations most people will feel with an older female caregiver.

An individual's response to touch can be influenced by a variety of other factors, including their sex, age, education, culture/language, socioeconomic status (SES), religion, and more (Remland, Jones, & Brinkman, 1995). For example, Eastern and Western culture has been shown to play a role in how pleasurable touching is perceived, with Westerners feeling more pleasure than Asian participants, but a similar role for emotional bond via social touch (Suvilehto et al., 2019). For this reason, we chose a Western-looking female actor. From this research we also learn that touch between adults is mediated by the relationship between the toucher and the touched. A closer relationship between the two is associated with more touch and more allowed areas of touch (Suvilehto et al., 2019). For non-Western participants, a human hugger from their own culture and ethnic identity would be more harmonious, but as this study was conducted in a metropolitan German city, we chose a Western female.

Additionally, research supports that women have a significantly stronger in-group bias than men (Rudman & Goodwin, 2004). These researchers also found a women have a pro-female bias that they automatically favored their mothers over their fathers or associated male gender with violence. This trend extends beyond just a preference; women are more comfortable hugging with other women than men are hugging other men. Some men are opposed to same-sex touching, and this fear appears to be of being perceived by others as homosexual (Floyd, 2000).

We also know that religion can play an important role in how people receive touch. Some studies highlight the challenges of providing medical care for Muslim women and men (Tackett et al., 2018). Some of the barriers for medical care, are similar for hugs. For example, Muslim women prefer a female healthcare provide whenever possible, or at least a female chaperone present throughout the encounter. Therefore, to ensure the majority of our participants felt comfortable interacting with our human hugging agent, we employed a female hugger.

While all the hugs in the experiment were platonic, we were also aware that people come with their own experiences. Some users may have a history of personal trauma, sexual abuse, or violence. Most physical violence in most cultures, is perpetuated by males against women (Watts & Zimmerman, 2002). For this reason, hugging an unknown male could be potentially very triggering for some users (Fosshage, 2000; Perry & Szalavitz, 2006; Statman-Weil, 2015). While women certainly can be the offender in situations of sexual assault or exploitation, we anticipated based on past research that by giving both our robot and female actor a "grandmotherly" appearance that users of both genders would feel equally comfortable hugging the agents.

As puberty approaches, neurohormones for survival of the species kick in and sexuality adds a new dimension to how users perceive touch. Based on research exploring visual stimuli and arousability among men, sexual connotations are given to a women with a greater waist to hip ratio, as it this indicates greater fecundity, and thus breedability (Singh, 1993). Therefore, we padded the actresses' waist to reduce her wait to hip ratio, giving her a more rotund appearance, rounded, grand-motherly appearance. This served a dual purpose, it helped our actress more closely match the appearance of the robot, which was also round, waistless/hipless and also, hopefully, decreased any sexual connotations our participants might have imagined.

7.3.3 *Behavior of Hugging Agents*

We trained the human actor to behave very similarly to the robot to ensure that the robotic human hug and the active robot hug were comparable. In the human hug condition, the actor behaved normally and much less choreographed. When she saw a user was ready for a hug, she slightly lifted and outstretched her arms to where it was comfortable and said, "I am ready for a hug." She closed her arms and moved as she normally would during a hug with friends or family.

To prepare for the robotic human hug, the actor experienced several hugs with the robot. She practiced with pilot participants before beginning the experiment, and she also reviewed videos of her hugs alongside robot hugs to catch any discrepancies. When the hugging agent detected a person, both the robot and the robotic human lifted their arms and said to the user, "I am ready for a hug." They waited with their arms outstretched until the user began walking toward them. At this point, they started closing their arms at a similar speed of around 20 cm/s.

In a passive robot hug condition, the robot closes its shoulder pan (joint 2), and elbow flex (joint 3) angles by only 15°, which was large enough that the arms did not touch any of our users. We consider this passive robot hug condition to be similar to an "air hug." This condition is meant to mimic other hugging robots created that users can hug, but the robot cannot adaptively hug them back.

For all conditions, the user controls the duration of the hug. The only exception to this is if the user does not initially apply enough pressure to the robot's back, in which case the robot will release after a set duration. More information on this special case is provided in Sub-Section 7.3.4. In all active hug conditions (human hug, robotic human hug, and robot hug), both the robot and the actor waited to release the user until they felt either the user remove their arms from their back or the user lean against their arms. At this point, they opened their arms and released the user.

In a passive robot hug condition, since the robot's arms are more open during the hug, the user was free to remove their arms and walk away at any point during the hug. While the torque release system was still working in the passive robot hug, no users contacted the robot's arms to activate this kind of release.

The robot and actor drop their arms when the user walks away from the hug in any condition.

7.3.4 *HuggieBot 4.0 Upgrades*

Compared to HuggieBot 3.0, we implemented several significant changes to improve the overall quality of the hug. First, we further refined the comfort and quality of the embrace by slightly adjusting the joint torques and joint angles of the closed pose position.

Next, we refined the visual perception of the robot to be more perceptive to user approach. Previously, the robot lifted its arms and invited a user for a hug, closing its arms when it detected forward movement from the user and lifting its arms based on the user approximated height. After the arm closing was activated, the visual perception of the robot stopped. If the user stopped walking forward, changed pace, or decided not to get a hug, the user would miss the hug, and the robot would hug without a user inside its arms. In past experiments, users were quite frustrated initially if they had difficulty getting the timing correct to be enveloped by the robot. In HuggieBot 4.0, we continuously monitor user movement. If the user stops walking forward or walks away, the robot reopens its arms to the invitation-for-a-hug pose.

Another change we implemented was handling when users do not initially apply enough pressure to the robot's back. This idea was inspired by Campbell and Yamane (2020), who also implemented solutions of edge cases for different kinds of unexpected user hugging behaviors. In our case, in past experiments, we noticed some users failed to apply the necessary amount of pressure on the robot's back to initiate a pressure release when they removed their hands. While we have dual hug release detection methods, not all users would remember to activate the torque release by pushing their backs against the robot's arms. Users sometimes became agitated, nervous, and panicked when the robot did not immediately release them. Therefore, we created an edge case that comes into effect at the beginning of the hug. If the robot's arms fully close and it has not detected an increase in pressure, it determines that the user did not apply sufficient pressure to the robot for it to detect a release, and thus, rather than waiting for the user indication to end the hug, it hugs on a timed schedule. Based on other research on human-human hugs, we determined that the hug would last three seconds (Nagy, 2011). From our previous experiment with HuggieBot 3.0, we also chose to have the robot perform a proactive squeeze before releasing, as we found in the previous chapter that most users appreciated this intra-hug gesture from the robot.

While in our previous study, we had the robot perform several different kinds of intra-hug gestures (rubs, pats, and squeezes), the Kinova JACO arms had some difficulties with the rubs and pats, which occasionally resulted in random and forceful arm jerks. After much investigation into the cause and discussion with engineers at the company, we were unable to predict when these jerks would occur or determine what caused them. Thus, for the safety of our users, we decided for this version of the robot to only have it perform squeezes on the user. The robot still detected and classified user rubs, pats, and squeezes. However, the behavioral algorithm always had the robot perform a squeeze in response. Future hugging robots should be able to respond safely to users in a variety of ways, rather than with only one gesture.

7.4 USER STUDY METHODS

This study aims to investigate the physiological effects of robot hugs on the human body. Following an adapted protocol from the Trier Social Stress Test, a standard method for ethically and reliably inducing psychological stress on participants in a laboratory setting (Birkett, 2011; Kirschbaum, Pirke, & Hellhammer, 1993), we measured participants' physiological responses to a human hug, a robotic human hug, a robot hug, a passive robot hug, and no hug. Figure 7.4 shows the timeline for the entire 3.5-hour duration of the study.

7.4.1 *Participants*

Fifty-seven participants were English-speaking volunteers recruited from Stuttgart, Germany. Five participants (four male, one female) were removed for either non-compliance or inappropriate behavior, leaving us with fifty-two users. No participants were employed by the Max Planck Society, so all were compensated at a rate of 10 euros per hour. Two participants had previously participated in an experiment featuring an earlier version of HuggieBot. For most users, it was their first robotic interaction of any kind.

We had a near-even gender balance (male = 51.92%, female = 48.08%). While our participants were mainly highly educated (92.3% of users holding

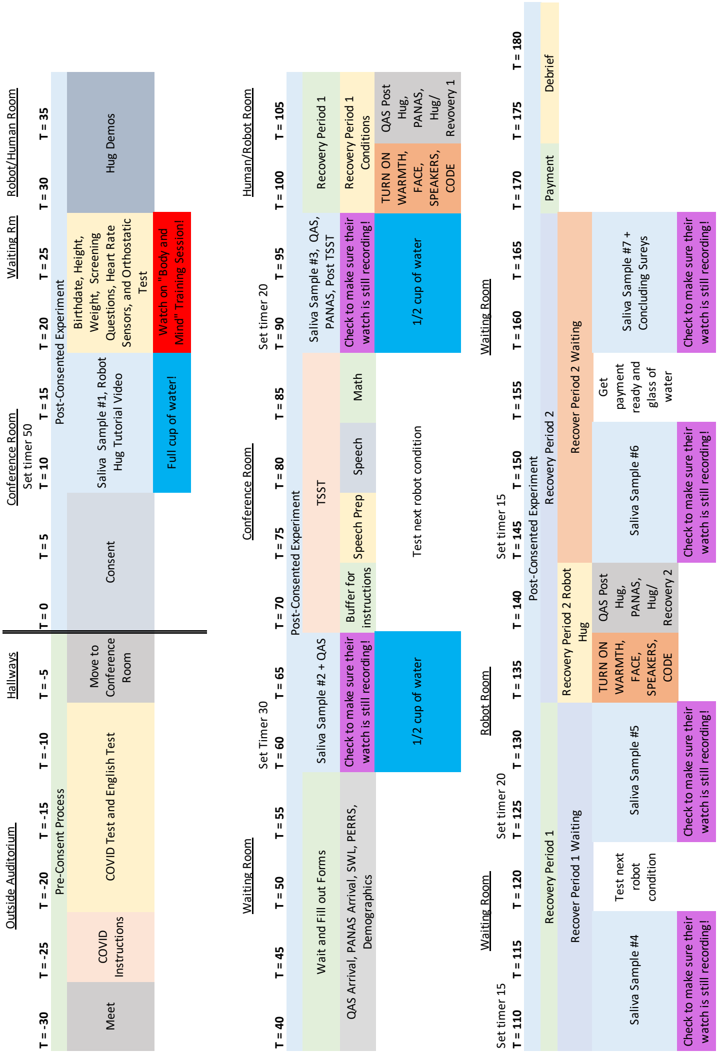


FIGURE 7.4: The timeline of the user study. The colored boxes represent the different phases of the experiment. To the left of the double-line represents the time before the participant has consented. QAS, PANAS, SWL, PERRS are questionnaires given to the participants, explained in the text of Section 7.4. TSST stands for the Trier Social Stress Test, the elements of which are explained below. Note that the hug provided in Recovery Period 1 Conditions is dictated by the user's predetermined hug group (including no hug), while all users receive an active robot hug in Recovery Period 2 Robot Hug.

a bachelor's degree or higher), they were primarily non-technical (86.54% self-identified as either "novice" or "beginner"). The average height of our users measured at the beginning of the experiment was 1.69 m, with a standard deviation of 0.10 m. Figure 7.7 shows the distribution of our users' age, weight, height, and BMI. The participants ranged in age from 22 to 58 (mean = 30.92, standard deviation = 7.75) and came from 20 different countries. Ten users self-reported to be introverted, while eleven claimed to be extroverted. Thirty participants felt they were equally introverted as extroverted, while only one user did not identify as either. Compared to the general public, our users were very optimistic about the future, as can be seen in Fig 7.8.

7.4.2 *Inclusion/Exclusion Criteria*

All users were required to be over 18 years old. All users were also required to speak and understand English well (roughly at B2 level, which was checked with an online exam upon arrival). Because of the large effect it can have on oxytocin levels, we excluded people who were pregnant, breast-feeding, or who had taken hormones in the three months before the study (such as oral contraceptives) when recruiting participants. Additionally, no one with uncorrected motor or vision disabilities was allowed to participate. Finally, no one with COVID-19 symptoms or recent exposure to someone with the symptoms or recent visit to a risk zone was able to participate in the study.

7.4.3 *Procedure*

After receiving approval from the Ethics Council of the Max Planck Society, we ran this study with additional measures to protect the health and safety of both the experimenters and the participants. The entire procedure follows a slightly modified Trier Social Stress Test (TSST) (Kirschbaum, Pirke, & Hellhammer, 1993).

7.4.3.1 *COVID-19 Health and Safety Measures*

Upon arrival, participants signed a document confirming that neither they nor anyone they are in contact with had recently visited a COVID-19 risk area, nor had they or anyone they were in contact with recently experienced any symptoms of COVID-19. Throughout the experiment, participants

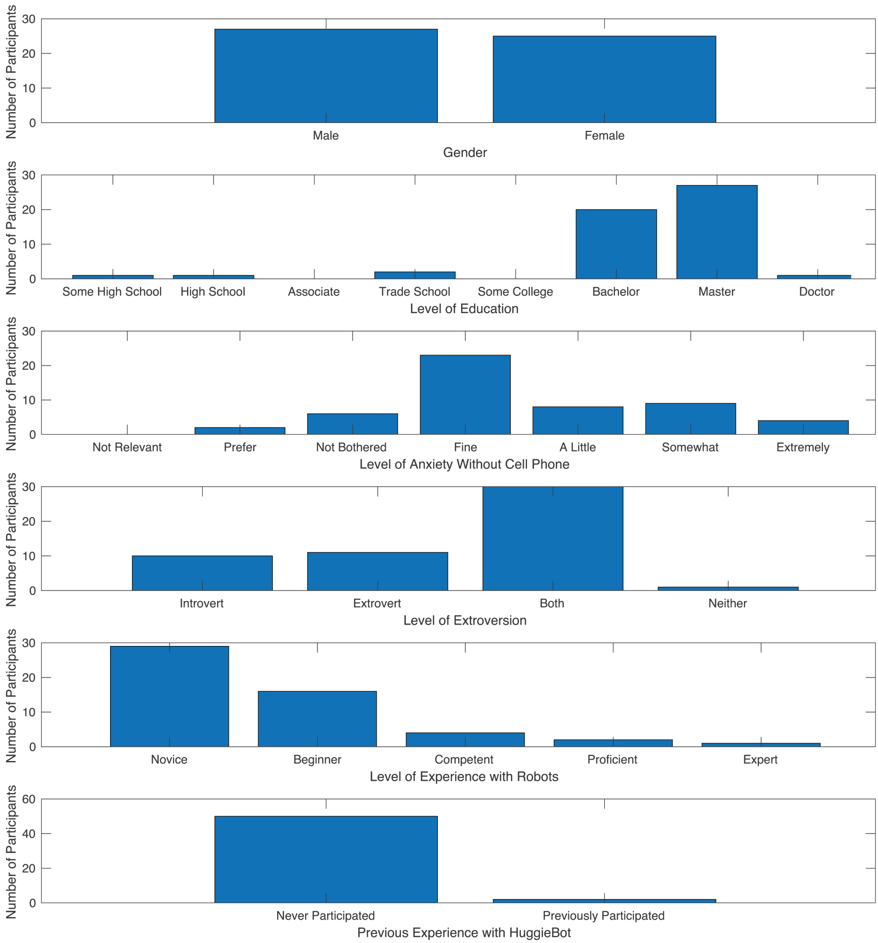


FIGURE 7.5: Bar graphs of the demographic distribution of the population of our user study based on gender, education, anxiety of being without a cell phone, extroversion, experience with robot, and prior experience with HuggieBot.

were asked to wash their hands thoroughly and/or use hand sanitizer. Whenever an experimenter and participant came with 1.5 m of one another, both people were required to wear an FFP2 mask. Each room used for the study had windows that remained open for proper ventilation and air circulation of the room. The robot and actor wore freshly laundered outfits for each participant. The outfits worn during the day were all washed at 60°C and thoroughly dried before they were re-worn. All common surfaces (chairs, tables, the robot head, heart rate monitors, chest straps, etc.) were wiped down and sanitized between participants. When handling participant saliva, the experimenter always wore clean, new disposable gloves. The saliva samples were stored in a locked freezer to prevent anyone from accidentally coming into contact with participant saliva.

7.4.3.2 *Pre-Consent*

All users were sent pre-visit instructions via email prior to their arrival. The instructions explained that they would be required to take and receive a negative result from a COVID-19 rapid test we provide and pass a B2-level English test to participate in the study. Potential users were also told that the entire experiment would be video recorded. To avoid contamination as a result of food intake participants were instructed to brush their teeth one and a half hours before they arrived (Rai et al., 2011; Salimetrics & SalivaBio, 2011). They were also told not to eat or drink anything other than water for one hour before their arrival to avoid affecting the oxytocin or cortisol levels detected in their saliva (Kirschbaum, Pirke, & Hellhammer, 1993; Rai et al., 2011; Salimetrics & SalivaBio, 2011). Finally, they were also requested to awaken a minimum of four hours before their appointment time and not to fall back asleep. We gave this sleep requirement because cortisol is known to spike the first few hours after rising (Brummett et al., 2009; Rimmele et al., 2010; Wilhelm et al., 2007).

Before participants arrive, they were split randomly into five groups corresponding to the five lettered conditions listed in Section 7.2. We took special care to ensure a minimum of five men and five women in each condition. Participants were met at the entrance of the building by a member of the experiment team. They were taken to a nearby seating area and asked to sign a consent form to take a self-administered COVID-19 rapid test. They were informed that if the test result was positive, they would be required to return home immediately and contact a medical professional. After they signed a COVID-19 rapid test consent form, we showed the participant a video demonstrating and explaining how to self-administer a COVID-19 rapid test. They were free to watch the video as many times as necessary. Additional written instructions were also included in the box. Users then took the test and either washed their hands or cleaned them using hand sanitizer (the choice was up to the participant).

As the self-test results took fifteen minutes to register, participants took a fifteen-minute online Cambridge Assessment General English test (https://www.cambridgeenglish.org/test-your-english/general-english/). Users who scored "intermediate" or "advanced" were welcomed to participate in the study. If a user scored "beginner", he/she did not meet the English requirement. All participants met the English requirement, and thus none had to be turned away. Additionally, all participants received a negative COVID-19 result, and thus none had to be turned away.

After receiving the English and COVID-19 rapid test results, users were asked if they still wished to participate. If they did, they were brought up to the fifth floor of the institute and taken into our Conference Room, where they met with the experimenter.

7.4.3.3 *Consent and Set Up Process*

Once in the Conference Room, users were reminded that the entire experiment would be video recorded. Users were given an optional media release consent form to sign. After participants indicated their approval or rejection of the media release consent form, the experimenter thoroughly explained the informed consent document. Users were given additional time to read both documents and encouraged to ask the experimenter any questions before signing anything.

Throughout the entire study, at seven different points in time, we collected saliva samples. The first sample was collected after the consent form was signed. The experimenter brought the participant an insulated cup filled with dry ice, two SaliCap sampling tubes, and a clean straw. She then explained the process and various techniques to provide saliva. Before the user began the first saliva sample, they watched a short 10-second video clip of someone rolling, cutting, and squeezing a lemon, to help aid saliva production. Participants were encouraged to lean their heads forward and allow saliva to pool to the front of their mouths. They were discouraged from swallowing during this process to help ease saliva production. Participants were free to either spit directly into the tube or use the straw to direct the saliva. The experimenter drew a blue line just below 1.5 ml to clearly identify to the participant the volume of saliva they were required to provide. Users were told that the liquid needed to reach the blue line and that bubbles could only begin above this line. The experimenter taught the participant several techniques to remove the bubbles from the saliva sample. Participants were told to hold the sampling tube by the cap, rather than by the base of the vial, to avoid applying heat to the saliva, which degrades the hormones. They were told to fill up both tubes at roughly the same time. They were encouraged to spit in one tube, place it in the dry ice (to help cool down the saliva immediately), spit into the other tube, and repeat this process until both tubes were completely filled. Then, users would show the tubes to the investigator, who would determine if they were adequately filled or if the participant needed to continue spitting.

The experimenter recorded the time (hour, minutes, and seconds) that the user began and finished providing each of the seven saliva samples. Once the user finished providing the first saliva sample, the experimenter gave him/her a glass of water to drink and had him/her watch a robot hug tutorial video, which gave instructions on how to properly hug the robot, while she put labels on the participant's saliva samples and immediately stored them in a freezer at $-80°C$ in a room just across the hall.

After the video concluded, the experimenter recorded the user's birthdate (day, month, and year). Using a doctor's scale, we recorded the participant's height and weight. Users were instructed to step onto the scale backward so they could not see their height or weight. Even when asked, the experimenter did not reveal the participant's height or weight until the conclusion of the experiment, in case these numbers could affect the user's emotional state. Then, users were asked to confirm that they fit the inclusion criteria by confirming they were not a) pregnant, b) breastfeeding, or c) taking any hormone medication.

Once confirmed, participants were shown to a changing area behind a privacy screen. They were given a key and shown a locker where they placed their personal belongings (including watches, cellphones, and any other potentially distracting items). They were also given a Polar H10 Heart Rate Monitor Chest Strap. A poster was placed on the wall next to the lockers, clearly identifying the correct and incorrect location/placement of the chest strap, which the experimenter also explained. Users were also given a small glass of water to wet the sensor before placing it directly against their skin. The experimenter left and closed the privacy screen around the user while he/she placed the chest strap under their clothes. The experimenter was available if the user had any questions or required any assistance. While the user stored their belongings and put on the chest

FIGURE 7.6: View of the couch and chair in the Waiting Room where the user performed the Orthostatic Test.

strap, the experimenter entered the user's gender, birthdate, height, and weight into the Polar V2 Vantage Watch.

When the user finished storing personal items and donning the chest strap, the experimenter asked for consent before holding the watch up close to the user's chest and pairing the watch to the chest strap. The experimenter then handed the watch to the user and asked him/her to fasten it snugly on their left wrist. The experimenter then led the user across the hall to the Waiting Room. The investigator explained that she would now like the user to perform an orthostatic test, which consisted of 90 seconds lying flat and still on a couch (in the Waiting Room) seen in Fig. 7.6, and then 90 seconds standing still. After the user asked any questions and agreed to perform the test, the experimenter asked for permission before touching the user's watch and initiating the test. The experimenter then left the room and informed the user that she would return before the end of the test. She also told the user that the watch would vibrate and the words "Stand up" would appear on the watch's face when it was time to change positions. When she left the room, the experimenter informed the actress playing the human hugging agent that the hug demonstrations would be beginning soon so that the actress could get ready in the correct costume, based on the user's pre-assigned condition.

Hug Demonstration Options
Human hug, then robot hug.
Human hug, then passive robot hug.
Robotic human hug, then robot hug.
Robotic human hug, then passive robot hug.
Robot hug, then human hug.
Passive robot hug, then human hug.
Robot hug, then robotic human hug.
Passive robot hug, then robotic human hug.

TABLE 7.1: The eight hug demonstration possibilities. We counter-balanced presentation order and ensured similar numbers of users experienced each of the eight options.

Once the orthostatic test was completed, the experimenter saved the result and began continuously recording the user's heart rate. If the orthostatic test failed (due to lost signal between the chest strap and watch), the user was brought to a sink, asked to apply slightly more water to the chest strap, and the test was restarted. Once continuous heart rate monitoring was initiated, the watch was locked, and the user was asked not to touch to watch for the remainder of the experiment to avoid accidentally stopping the recording. Then, the user was taken to see live hug demonstrations with both the human hugging agent and the robot hugging agent. The order of the demonstration was randomized between the eight options shown in Table 7.1.

Each user saw both a human hug and a robot hug demonstration. The presentation order between human and robot agents was counterbalanced. Unless the user was assigned to the "no hug" condition for the first intervention, one of the two demonstrations was the first intervention he/she would receive (though he/she was not informed of this). In each given condition, we counterbalanced so that roughly half the women and men had a robot demonstration first and half had a human demonstration first.

After the hug demonstrations, the user was brought back to the Waiting Room, given a clipboard and a clean pen, and asked to fill out several surveys while the experimenter "prepared the rest of the experiment." The users were told to do their best on all the surveys and to circle any questions they had, and to move on. The experimenter promised to come back in

Arrival Survey Order
QAS (Quantitative Affect Scale) Arrival
PANAS (Positive and Negative Affect Schedule) Arrival
SWL (Satisfaction With Life - Table 7.3)
PERRS (Personal Empathy Response and Regulation Survey - Table 7.4)
Demographics Questionnaire

Table 7.2: The five surveys users were asked to fill out during the waiting period. At this point, users had arrived to the study, taken a COVID-19 test and an English comprehension test, consented to participate, provided the first saliva sample, started the heart rate recording, and seen both hug demonstrations.

Satisfaction with Life - Questions
In most ways, my life is close to ideal
The conditions of my life are excellent
I am satisfied with my life
So far I have gotten the important things I want out of life
If I could live my life over, I would change almost nothing

Table 7.3: The five survey questions asked in the Satisfaction with Life survey. Users were asked to rate how much they agreed with the following statements on the given scale (1 - not true at all to 7 - absolutely true)

ten minutes to answer any questions they had. If they finished all the survey questions before the experimenter returned, the users were told to sit quietly and rest. The experimenter had a timer, and each user spent twenty minutes in the Waiting Room after the hug demonstrations. During this waiting time, the users filled out the surveys in the order seen in Table 7.2.

Some participants had difficulty finishing the Demographics questionnaire within the given time. They were allowed to hold onto this form and fill it out at their leisure throughout the experiment whenever they had time.

After twenty minutes, the experimenter returned to the Waiting Room with the insulated cup filled with dry ice, two empty sampling tubes, and a clean straw for the participant. She collected the completed survey

Personal Empathy Response and Regulation Survey - Questions
I usually understand what is going on between people in emotional situations
I usually understand and label other people's feelings correctly
I often feel the same feeling that I see someone else feeling
When someone is getting hurt or feeling strong feelings, I usually feel something like they feel in my body, too (tingling, pain, upset stomach, dizzy, etc.)

Table 7.4: The four survey questions asked in the Personal Empathy Response and Regulation Survey. Users were asked to rate how much they agreed with the following statements on the given scale (1 - totally disagree to 9 - totally agree)

forms and gave the saliva collection materials to the participant. Then, the experimenter went to the back of the room to record the time the participant started and finished providing saliva. Once the user closed the tubes and handed the cup of dry ice to the experimenter, she checked the participant's watch to ensure it was still recording. The experimenter gave the participant a half cup of water and another QAS survey (Post Wait) to fill out while she labeled and put the saliva samples in the freezer.

7.4.3.4 *Trier Social Stress Test (TSST)*

The experimenter then collected the QAS Post Wait survey and the empty water cup and led the participant back to the Conference Room to begin the Trier Social Stress Test (TSST), which is composed of three parts: Speech Prep, Speech Task, and a Math Task. This combination of public speaking and oral mathematics has been shown to ethically and reliably induce stress in participants (Birkett, 2011; Kirschbaum, Pirke, & Hellhammer, 1993).

Speech Prep

The user was given a piece of paper and a pen and instructed that he/she had five minutes to prepare a five-minute speech about why he/she is the ideal candidate for their dream job. He/she was told that he/she could take notes and help structure their talk, but that he/she would not be able to use the notes during the presentation. The user was informed that the speech would be recorded and performed in front of a panel of experts. The experimenter set a timer for five minutes and left the participant alone in the room to prepare.

Panel of Experts

After five minutes, the experimenter returned to the room with the panel of experts.

The Panel of Experts always consisted of three members, with at least one male member and one female member. Given scheduling constraints, four actors (two male and two female) rotated through the roles of the Panel of Experts. All members of the Panel of Experts wore white lab coats and carried clipboards, notebooks, and pens.

When they entered the Conference Room, the experimenter introduced them as "an expert in public speaking, an expert in detecting the difference between true and false statements, and an expert in mathematics."

The participant's notes were taken away, and the user was then asked to stand on a tape "x" marked on the floor at the end of the long conference table. Due to the length of the table and the open windows for ventilation, the participant was invited to remove their mask. Participants only removed their masks if they felt comfortable doing so, and it was only for this part of the experimenter. The Panel of Experts kept their masks on the entire time because they sat close together.

The experimenter then explained to the participant that she would be leaving because she "is not a part of this task of the experiment," but that she would return to collect the participant after the task. She informed the participant that the Panel of Experts would give him/her all further instructions needed.

All the members of the Panel of Experts were instructed to observe the participant throughout the stress test and take notes on their behavior. They were also given evaluation forms to evaluate the participant's performance in the Speech and Mask Tasks.

Speech Task

For the Speech Task, the Public Speaking Expert said to the user "This is the speech portion of the task. You are to deliver a speech describing why you would be a good candidate for your ideal job. You should speak for the entire the five-minute time period. Your time begins now."

At this point, the participant began to give their speech. If the participant stopped talking during the speech, the Panel of Experts remained silent while staring at the participant for twenty seconds. If he or she did not resume speaking, the Public Speaking Expert prompted the participant to continue speaking by instructing him/her, "You still have time remaining."

If the participant stopped again, the Public Speaking Expert reminded him/her one more time after waiting quietly for twenty seconds that there was time remaining.

If the participant stopped a third time, the Panel of Experts sat in silence staring at the participant until the five-minute time period had elapsed.

Math Task

As soon as the timer went off that the five-minute Speech Task was over, they moved directly to the Mask Task. The Public Speaking Expert instructed the user, "During the final five-minute math portion of this task, you will be asked to subtract the number 13 from 1,022 sequentially. You will verbally report your answers aloud and be asked to start over from 1,022 if a mistake is made. Your time begins now."

The participant then begins to report the answers verbally. If the participant makes a mistake, the Mathematics Expert hits a buzzer, making a loud noise and a red "x" lights up. The Mathematics Expert says, "That is incorrect. Start over from 1,022." The user must then start over from the beginning until the entire five minutes have elapsed.

The Panel of Experts kept track of how many times each user made a mistake and had to start over and how far they reached before restarting.

Post TSST

After the conclusion of the Math Task, the experimenter returned to the Conference Room, and the Panel of Experts left. The experimenter did not speak with the participant about the tasks he/she just experienced, even if the participant wanted to. Immediately, the user was asked to provide another saliva sample (the third). The user was given the saliva collection materials, and the experimenter recorded the spitting start and end times.

While the experimenter took the saliva samples to be labeled and placed in the freezer, the user was given the last half cup of water allowed for the experiment and three surveys: QAS Post TSST, PANAS Post TSST, and a Post TSST Reflection survey.

7.4.3.5 *Recovery Period 1*

Recovery Period 1 comprises a ten-minute intervention session to experience the assigned hugging condition, the fourth saliva sample, a waiting period, and the fifth saliva sample, as described below:

Intervention 1 - Conditions

Once the user completed all three surveys, the experimenter took him/her to either the Human Room or the Robot Room, depending on the condition to which he/she was assigned.

If the user was assigned to any of the hug conditions, the experimenter's instructions were the same: "You will now have ten minutes in this room. During this time, we would like to request that you perform a minimum of one hug, but there is no maximum number of hugs. We are interested in studying the effect of touch on neuro-hormones, so we encourage you to hug for the maximum amount of time that is comfortable for you. To clarify, you can hug the agent as many times as you want for as long as you want. The choice is always up to you. When you are ready, there are a few survey questions on this clipboard for you that I will place on the side table. Please complete at least one hug before filling out the surveys. You may answer the questions and then receive more hugs, and if those interactions change your mind, you are welcome to change your responses to the questionnaire. You may not leave the room before the time is up, but how you choose to spend your time in this room is completely up to you. If you no longer wish to hug, you are welcome to have a seat on this stool at any point. Your time begins now. Please wait for the agent to let you know she is ready."

If the user was assigned to the no hug condition, the instructions the experimenter gave were slightly different, as there was no hugging agent in the room with the user. In this case, she said, "You will now have ten minutes in this room. Please spend some time in the space however you like. When you are ready, after you have spent at least a few minutes in the room, there are a few survey questions on this clipboard for you that I will place on the side table. You may answer the questions, and at any point, if you change your mind, you are welcome to change your responses to the questionnaire. You may not leave the room before the time is up, but how you choose to spend your time in this room is completely up to you. If you no longer wish to explore the space, you are welcome to have a seat on this stool at any point. Your time begins now."

The experimenter then placed the clipboard with survey questions on the side table and sat at a table behind the hugging agent. During the ten minutes in the room, the experimenter looked down at her notebook and took notes. She observed the participant for inappropriate behavior and the safety of both the participant and the actress/robot.

If the user tried to speak with the actress in either the human hug or robotic human hug conditions, the experimenter said, "I'll be more than happy to answer any questions you have. The hugging agent can't respond to you right now." Since the robot cannot respond to users, we told the actor not to respond to participants to make the human and robot conditions more comparable.

If the user tried to speak with the experimenter in any conditions, the experimenter said, "I'll be more than happy to answer any questions you have," and answered the question. If the user tried to engage in conversation with the experimenter, she told the user "Please focus on your task. I'm just here for your safety and to answer any questions you have about the task given to you."

At the end of the ten minutes, the experimenter said, "Your time in this room is now up. Please bring the surveys with you and follow me back to the Waiting Room. If you haven't finished them during these ten minutes, that's no problem. You will have time to continue working on them in the Waiting Room."

Waiting 1

Once in the Waiting Room, the experimenter brought the user the saliva collection materials and asked him/her to provide the fourth saliva sample of the experiment. If the user was finishing up the surveys, the experimenter asked him/her to set them aside briefly, focus on providing the saliva sample, and go back to surveys once finished. The experimenter then went to the back of the room and recorded the time the user started and finished providing the saliva sample. The experimenter collected the closed tubes and the cup, labeled the samples, and placed them in the freezer. Before she left, she checked the user's watch to ensure it was still properly recording. The experimenter told the user she would put the saliva samples in the freezer and return in a few minutes. The experimenter told the user that while she was gone, he/she could use the time to complete any surveys he/she had not yet finished and otherwise to just rest and relax until the experimenter returned.

The experimenter came back with the fifth saliva sample collection materials fifteen minutes after the start of the fourth saliva sample (she used a digital timer to help ensure correct timing). The same process was repeated of bringing the saliva collection materials and recording the start and finish time of providing the saliva.

Two female participants took over twenty minutes to provide their fourth saliva sample, and thus the fifth saliva sample was omitted for them due to time constraints.

7.4.3.6 *Recovery Period 2*

Just as with Recovery Period 1, Recovery Period 2 is composed of a ten-minute intervention session. Regardless of their assigned condition, everyone experienced the robot hug this time; then came the sixth saliva sample, a waiting period, and the final seventh saliva sample.

Intervention 2 - Robot Hug

After labeling and putting the fifth saliva sample in the freezer, the experimenter took all users to the Robot Room.

The instructions the experimenter gave were the same as they were for the first intervention hug conditions: "You will now have 10 minutes in this room. During this time, we would like to request that you perform a minimum of 1 hug, but there is no maximum number of hugs. We are interested in studying the effect of touch on neuro-hormones, so we encourage you to hug for the maximum amount of time that is comfortable for you. To clarify, you can hug the agent as many times as you want for as long as you want. The choice is always up to you. When you are ready, there are a few survey questions on this clipboard for you that I will place on the side table. Please complete at least one hug before filling out the surveys. You may answer the questions and then receive more hugs, and if those interactions change your mind, you are welcome to change your responses to the questionnaire. You may not leave the room before the time is up, but how you choose to spend your time in this room is completely up to you. If you no longer wish to hug, you are welcome to have a seat on this stool at any point. Your time begins now. Please wait for the agent to let you know she is ready."

The experimenter then placed the clipboard with survey questions on the side table and sat at a table behind the hugging agent. During the ten minutes in the room, the experimenter looked down at her notebook and took notes. She observed the participant carefully for inappropriate behavior and the safety of both the participant and the robot.

At the end of the ten minutes, the experimenter said "Your time in this room is now up. Please bring the surveys with you and follow me back to the Waiting Room. If you haven't finished them during this ten-minute period, that's no problem. You will have time to continue working on them in the Waiting Room."

Waiting 2

Once in the Waiting Room, the experimenter brought the user the saliva collection materials and asked him/her to provide the sixth saliva sample of the experiment. If the user was finishing up the surveys, the experimenter asked him/her to set them aside briefly, focus on providing the saliva sample, and go back to surveys once he/she finished. The experimenter then went to the back of the room and recorded the time the user started and finished providing the saliva sample. The experimenter collected the closed tubes and the cup, labeled the samples, and placed them in the freezer. Before she left, she checked the user's watch to ensure it was still properly recording. The experimenter told the user she would put the saliva samples in the freezer and return in a few minutes. The experimenter told the user that while she was gone, he/she could use the time to complete any surveys he/she had not yet finished and otherwise to just rest and relax until the experimenter returned.

The experimenter came back with the seventh and final saliva sample collection materials fifteen minutes after the start of the sixth saliva sample (she used a digital timer to help ensure correct timing). The same process was repeated of bringing the saliva collection materials and recording the start and finish time of providing the saliva.

The same two female participants took over twenty minutes to provide this saliva sample, and thus the seventh saliva sample was also omitted for them due to time constraints. Consequently, they each gave only five saliva samples, while all other participants gave seven.

After the user finished providing the last saliva sample, the experimenter provided him/her with a glass of water to drink if he/she wished. She also gave the user the final closing surveys, which included a QAS, a PANAS, and a Concluding Reflection survey, which asked about the user's overall experience in the experiment.

7.4.3.7 *Payment and Debrief*

Once the user completed the surveys and handed them to the experimenter, she stopped the recording on the watch. The experimenter then gave the participant 35 euros for their participation and had him/her sign a document confirming that he/she received the payment. At this point, the experimenter also gave the user two additional documents. The first was a letter written by the experiment team explaining we understand the user went through a challenging experience. We wrote that everyone finds the public speaking and verbal math tasks challenging and that the user performed normally. We apologized for the inconvenience and anxiety they felt, and we thanked them for undergoing this experience because it will be beneficial to our research. Many users found this letter to be extremely comforting. A final document we provided to the user was a two-page sheet explaining several different tips, tricks, and techniques on alleviating stress and anxiety in their daily lives.

The experimenter explained in a little more detail about the experiment and asked the user if he/she had any additional questions, which she would be happy to discuss now. After the participant felt all their questions were answered, they returned to the Conference Room. The participant removed the Polar H10 Chest Strap and Vantage V2 Watch and handed them to the experimenter. The user collected their items from the locker and was excused from the experiment.

7.4.4 *Salivary Analysis*

As previously mentioned, all participants' saliva samples were stored in a freezer in our lab at $-80°C$ until they were ready for transport. Then, we relocated them to the Institute for Medical Psychology at the University of Heidelberg for analysis. During the transport, the saliva samples were stored in tightly sealed styrofoam boxes filled with dry ice. The entire trip took about an hour and a half, and the samples were still frozen upon arrival.

The lab analyzing our samples for cortisol used a Demeditec ELISA Kit. The detection range of the assay used for the cortisol analysis is $0.1 - 30$ ng/ml. The sensitivity of the Demeditec assay is 0.019 ng/ml.

The lab conducting the oxytocin analysis of our saliva samples used an Enzo ELISA Kit. The detection range of the assay used for the oxytocin analysis is $15.6 - 1,000$ pg/ml. The sensitivity of the Enzo assay is 15 pg/ml.

7.5 RESULTS

Throughout this study, we collected several different forms of data, which we present here and discuss in Section 7.6. We present data on our demographics and participant pool, the many different survey responses (emotional measures), the robot data, and several physiological measures.

7.5.1 *Demographics and Participant Pool*

In this subsection of the results, we discuss the demographics of our participant pool. We include the collected experimenter information (like age, height, and weight) and their responses to the Gallup poll. We asked our users about their frequency and enjoyment of hugs. We also had them fill out a Satisfaction with Life survey and the Personal Empathy Response and Regulation Survey. Together, all this information gives us an insight into how our participants felt upon entering the research study.

7.5.1.1 *Collected Experimenter Information*

After the user provided the first saliva sample, the experimenter collected the user's birthdate, height, and weight. Figure 7.7 shows the birthdate converted to age, weight, height, and BMI (calculated from height and weight). Our participants were in the average range for height, weight, and BMI for the general population. Their average age was 30.92 years, with a standard deviation of 7.75. Our users were healthy, with an average user weight of 71.79 kg and a standard deviation of 14.81. Our average user height was 1.69 m with a standard deviation of 0.09. Finally, the average user BMI was 25.05, with a standard deviation of 4.12.

7.5.1.2 *Gallup Poll*

As part of our demographics questionnaire, we asked users three questions obtain some insight into the mindset in which they arrived at our study. We had our users fill out a Gallup Poll, which asks how users feel about the quality of their current and predicted future life (five years from now). Using a Wilcoxon-Signed Rank test, we found that our participants were significantly more optimistic about the quality of their future life compared to the state of their current life ($p < 0.001$), as can be seen in Fig. 7.8.

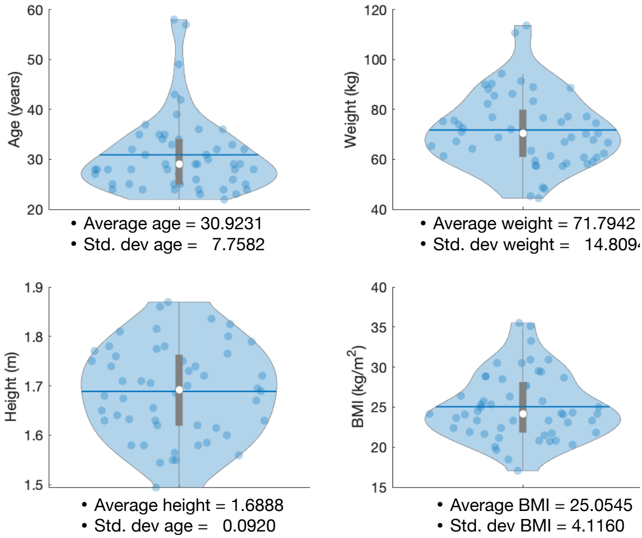


FIGURE 7.7: Violin plots showing the distribution of our participants' age, weight, height, and BMI. The shape of each violin represents the kernel density estimate of the data. Small dots are overlaid showing the exact data points. The white dot in the center represents the median, and the the bottom and top edges of the thick grey bar indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. The mean is shown by a horizontal line across the entire violin in the same color as the individual data points.

7.5.1.3 *Frequency and Enjoyment of Hugs*

We asked users how frequently and how much they enjoyed hugs before and during the COVID-19 pandemic. Each question asked users to answer for romantic partners, non-romantic family members, friends, and pets. Figure 7.9 shows how the users responded to these questions.

On average, our users highly enjoyed hugs from romantic partners, non-romantic family members, and friends. We also found that our users were significantly deprived of hugs! Our users had a significant decrease in hugs from romantic partners ($p = 0.0029$), non-romantic family members

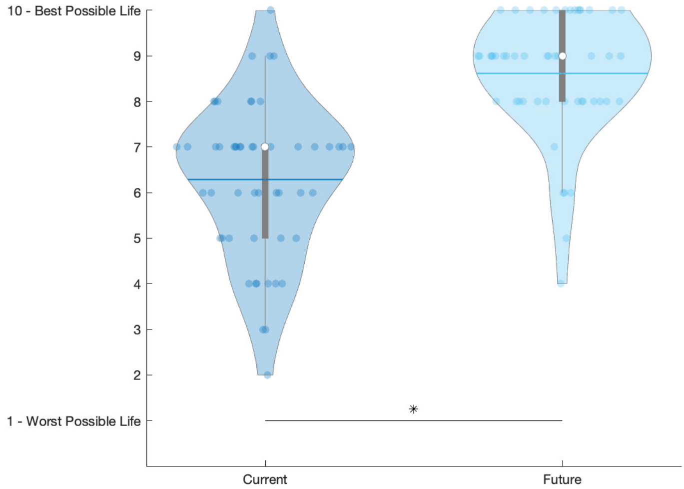


Figure 7.8: A violin plot of our users' responses to the Gallup Poll questionnaire, asking about the quality of their current and projected future life.

($p < 0.001$), friends ($p < 0.001$), and pets ($p = 0.0042$) in the last seven days compared to before the COVID-19 pandemic began.

7.5.1.4 *Satisfaction with Life (SWL)*

We also had participants fill out the Satisfaction with Life survey and the Personal Empathy Response and Regulation Survey (PERRS), which are validated psychology questionnaires to understand the mentality of participants entering a study. Figure 7.10 shows how our participants responded to the five questions asked in the Satisfaction with Life survey. Overall our participants were moderately satisfied with the current state of their lives. By combining all five questions, we find our users are moderately satisfied with their lives as they have an average satisfaction score of 4.2577 (on a scale from 1 - 7), with a standard deviation of 1.7127. No correlation was observed between self-reported satisfaction with life and basal oxytocin level. Using a Pearson's correlation, we found that user's self-reported satisfaction with life had a moderate, positive, linear correlation to their self-reported responses to the first question of the Gallup Poll, which asked about their current life quality ($\rho = 0.5103$, $p = 0.0001$).

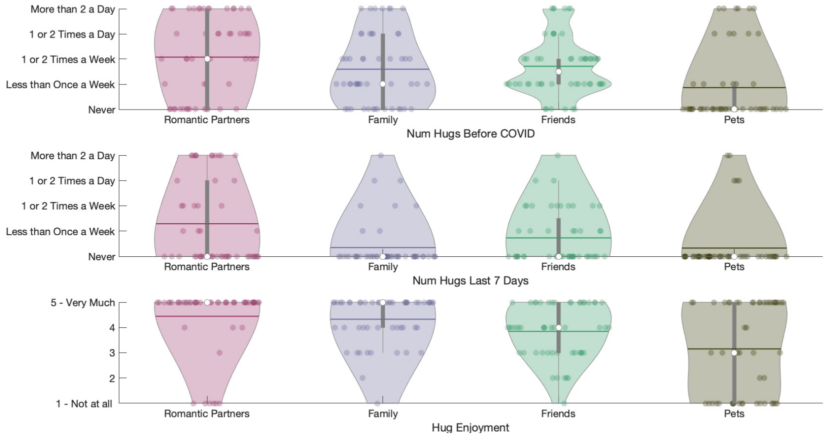


FIGURE 7.9: Violin plots showing user responses to the frequency and enjoyment of hugs in daily life separated by different types of hugging partners.

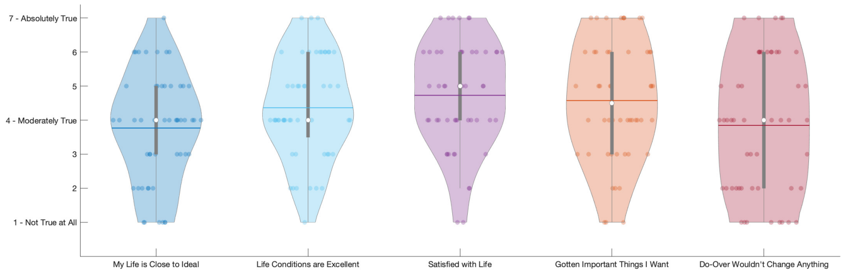


FIGURE 7.10: Violin plots showing the participant responses to all five questions of the Satisfaction with Life survey.

7.5.1.5 *Personal Empathy Response and Regulation Survey (PERRS)*

Figure 7.11 shows how our users responded to each of the four questions in the PERRS. Figure 7.12 combines the first three questions of the PERRS into a single value, showing the total empathic response. Since the data are non-parametric, we used a Wilcoxon-Signed Rank comparison test and found that our users had significantly more empathic responses than emotional arousal ($p = 0.0039$).

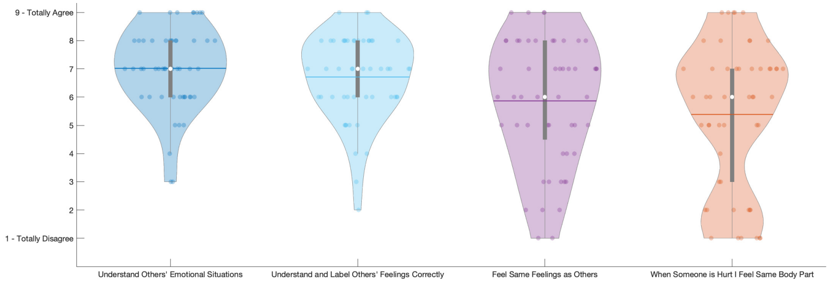


FIGURE 7.11: Violin plots showing the participant response to all four questions of the PERRS questionnaire.

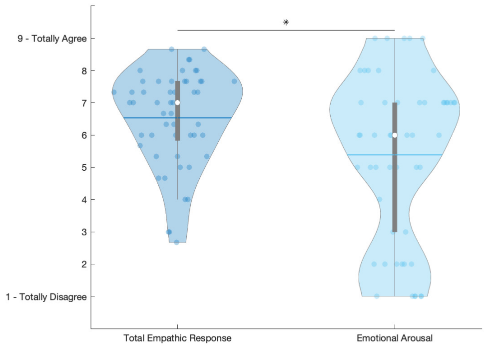


FIGURE 7.12: Violin plots showing the first three questions of the PERRS questionnaire combined to yield a "total empathic response" compared to the final PERRS question, which is renamed "emotional arousal."

7.5.2 *Survey Responses*

Throughout the experiment, users filled out several different surveys, self-reporting their emotional state and how they felt about different aspects of the study. This subsection will discuss the results from the Quantitative Affect Schedule (QAS), the Positive and Negative Affect Schedule (PANAS), the post-intervention ratings, the ratings to re-experience the intervention method, and the users' self-reported most beneficial session.

7.5.2.1 *Quantitative Affect Schedule (QAS)*

At six different time points throughout the experiment (upon arrival, after the waiting period, after the Trier Social Stress Test (TSST), after the first hug intervention, after the second hug intervention, and at the end of the experiment before the debrief), we asked users to fill out the Quantitative Affect Schedule (QAS). This survey has two parts: the first part asks users to quantify the emotional intensity they are currently feeling, the second part asks them to assess the quality of how they are feeling. The emotional intensity part asks the simple questions, "how are you feeling right now," and is broken into positive and negative emotions. The negative scale ranges from -5: extremely negative, -4: very negative, -3: negative, -2: somewhat negative, -1: a little negative, 0: neutral, to 0: not applicable. The positive scale ranges from 0: not applicable, 0 = neutral, +1: a little positive, +2: somewhat positive, +3: positive, +4: very positive, +5: extremely positive. Users were also given free spaces to explain why they felt negative and why they felt positive. These questions were not mutually exclusive; users could report both negative and positive emotions simultaneously.

QAS - Intensity: Figure 7.13, shows user responses to the emotional intensity questions across the six different time points we asked during the study, without separating them by the first hug intervention they received. Users felt significantly more intense negative feelings immediately after the Trier Social Stress Test (TSST) than at any other point in the experiment. To analyze this, we used a repeated-measures analysis of variance (`ranova` in MATLAB 2019b) with a significance value of $\alpha = 0.05$.

QAS - Negative: Since Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon = 0.66129$), we report the Greenhouse-Geisser corrected results. Overall, there was a significant change in users' self-reported negative emotional intensity throughout the six time points taken ($F(5, 115) = 10.012, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test, we found that from when users arrive to the experiment to immediately after the TSST, the intensity of their negative emotions significantly increases ($p < 0.001$), as it does from the post wait to immediately after the TSST, ($p < 0.001$). After the hug intervention, overall, the intensity of user negative emotions significantly decreases ($p = 0.0106$). There is also a significant decrease in negative emotional intensity between after the TSST to after the second hug intervention ($p = 0.00304$), and between after the TSST and the end of the experiment ($p < 0.001$). Users came in with slightly more intense negative emotions than they had after sitting in the waiting room, but by the end of the experiment, the intensity of their negative emotions was almost the same as it was after the short waiting period ($p = 0.99946$).

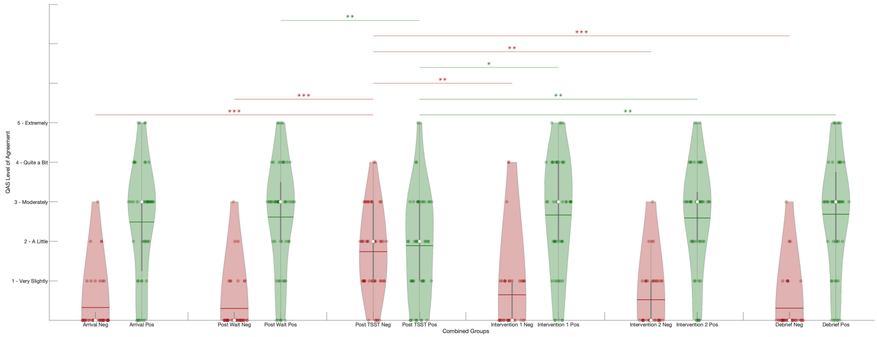


FIGURE 7.13: Violin plots of the participants' responses to the QAS questionnaire over the course of the study. The responses from all participants are shown together, regardless of the hug condition they were separated into. The absolute values of the negative emotions are shown in red, and the positive emotions are shown in green.

QAS - Positive: Users positive emotions also changed the most in comparison to after the TSST. Using Mauchly's test, we found that the assumption of sphericity had been violated ($\epsilon = 0.80146$). Thus, we report Greenhouse-Geisser corrected results. Overall, there was a significant change in users' self-reported positive emotional intensity throughout the six time points taken, ($F(5, 185) = 3.6403, p = 0.0036431$). We again used a Tukey-Kramer posthoc multiple comparison test and found a significant decrease in the positive emotional intensity been the post wait and immediately after the TSST ($p = 0.014875$). We notice an increase in the positive emotional intensity from immediately after the TSST to the first intervention ($p = 0.0508$). The second intervention also sees an increase in positive emotional intensity compared to after the TSST ($p = 0.00986$), as does the debrief at the end of the experiment compared to following the TSST ($p = 0.0139$).

QAS - By Condition: We were also interested in seeing whether the intervention users received immediately following the Trier Social Stress Test (TSST) affected their self-reported emotional intensity of positive or negative emotions. Figure 7.14 shows the positive and negative QAS responses of users at each time point collected in the study, separated by which first intervention method they received. We again used a Tukey-Kramer posthoc multiple comparison test following the same repeated-measures analysis of variance. This time rather than just looking at the within-subject factor of time, we also included the between-subject factor of hug condition.

No Hug: In the "no hug" condition, the increase in intensity of negative emotions approached significance from the arrival to after the TSST ($p = 0.077281$), and the post wait to after the TSST ($p = 0.077659$). No significant differences in the reported intensity of positive emotions were observed at any time points collected.

Passive Robot Hug: In the "passive robot hug" condition, no significant differences in the reported intensity of negative emotions were observed at any of the time points collected. Additionally, no significant differences in the reported intensity of positive emotions were observed at any time points collected.

Robot Hug: In the "robot hug" condition, there was a significant increase in the intensity of negative emotions from the arrival to after the TSST ($p = 0.0081545$), and the post wait to after the TSST ($p = 0.00127$). A significant decrease in the intensity of negative emotions was observed from the TSST to the second intervention ($p = 0.026889$, and the TSST to the debrief period ($p = 0.02229$). No significant differences in the reported intensity of positive emotions were observed at any time points collected.

Robotic Human Hug: In the "robotic human hug" condition, no significant changes in the self-reported intensity of negative emotions at any of the time points collected. A significant decrease in positive emotions was observed between the post wait and after the TSST ($p = 0.047208$). An increase in the intensity of positive emotions approaches significance between the TSST and the first intervention ($p = 0.071927$).

Human Hug: In the "human hug" condition, there was a significant increase in the intensity of negative emotions from the arrival to the TSST ($p = 0.029227$), and the post wait to the TSST ($p = 0.029408$). The decrease in intensity of negative emotions approached significance from after the TSST to the debrief period ($p = 0.065319$). No significant differences in the reported intensity of positive emotions were observed at any time points collected.

QAS - Quality: As previously mentioned, the second part of the QAS asked users to assess the quality of how they are feeling. At each time point the QAS was administered, users were asked to evaluate the same four listed emotions: cared for, loved, seen, and in sync. All four emotions were analyzed using a repeated-measures analysis of variance (`ranova` in MATLAB 2019b) with a significance value of $\alpha = 0.05$.

QAS - Cared For: Figure 7.15 shows the user responses to how "cared for" they felt at the six different time points across the study, separated by which hug condition they received during the first intervention. Using Mauchly's

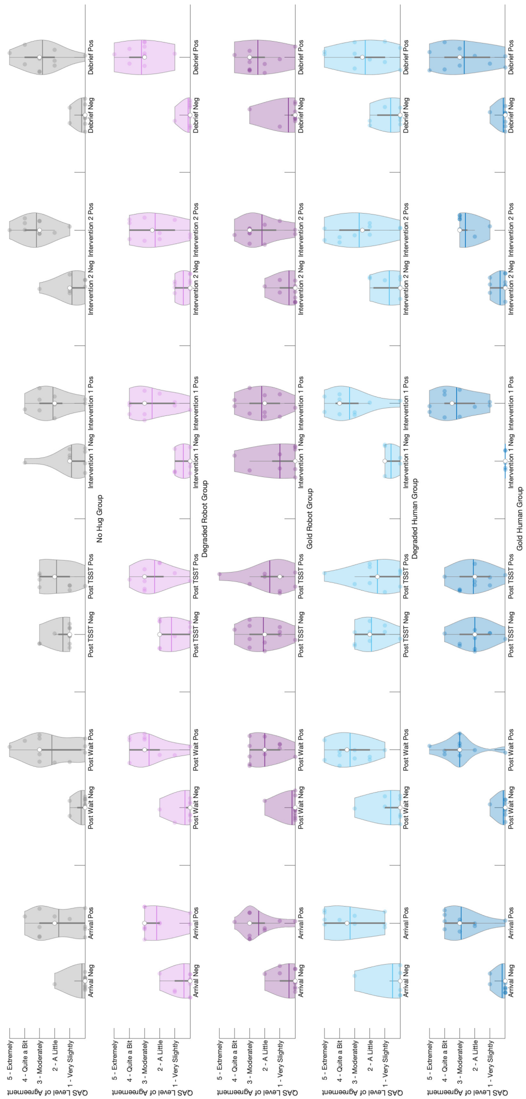


Figure 7.14: Violin plots of the participants' responses to the QAS questionnaire over the course of the study. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

test, we found the assumption of sphericity had been violated ($\epsilon = 0.67126$), so the Greenhouse-Geisser corrected results are reported. Overall, there was a significant change in how "cared for" users felt over the course of the experiment ($F(5, 235) = 7.7853, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found that users in the "no hug" condition felt significantly less cared for from the post wait to after the TSST ($p = 0.015088$), and from the post wait to after the first intervention, where they were alone in an empty room ($p = 0.033103$). Users felt significantly more cared for after the second intervention compared to how they felt after the TSST ($p = 0.029475$). Users in the passive robot hug group did not feel any significant changes in how cared for they felt throughout the study. The increase in how cared for people in the robot hug condition felt during the debrief compared to how they felt after the TSST approached significance ($p = 0.07733$). Users in the robotic human hug group did not feel any significant changes in how cared for they felt over the course of the study. Users in the human hug condition felt significantly more cared for after the first hug intervention compared to how they felt after the TSST ($p = 0.032865$).

QAS - Loved: Figure 7.16 shows the user responses to how "loved" they felt at the six different time points across the study, separated by which hug condition they received during the first intervention. Since Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon = 0.70436$), Greenhouse-Geisser corrected results are reported. Overall, there was a significant change in "loved" for users felt over the course of the experiment ($F(5, 235) = 7.9296, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found that users in the no hug condition felt significantly less loved after the first intervention compared to how they felt upon arrival ($p = 0.033913$). The decrease in feeling loved after the first intervention compared to how loved they felt after the waiting period approaches significance ($p = 0.070153$). Neither users in the passive robot hug condition nor in the robot hug condition groups experienced any significant changes in how loved they felt throughout the experiment. The robotic human hug condition group experienced a decrease in feeling loved that approached significance after the second intervention compared to the first intervention ($p = 0.079634$). Users in the human hug condition felt significantly more loved after the first intervention period than how they felt after the TSST ($p = 0.049364$).

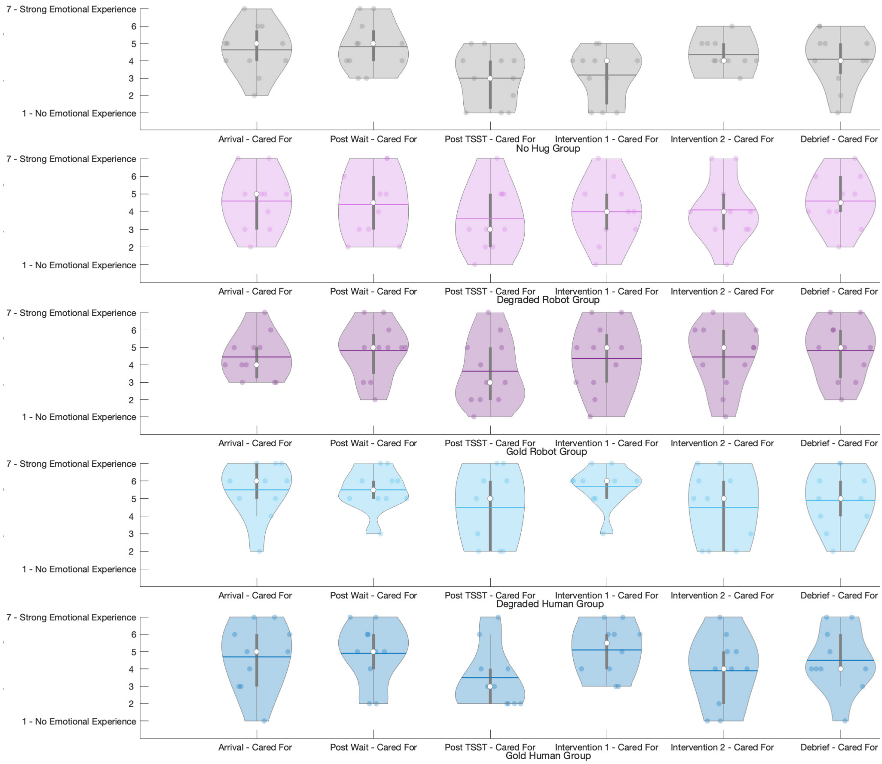


FIGURE 7.15: Violin plots of the participants' responses to the QAS question asking users to assess the quality of how "cared for" they felt at each of the six time points the data was collected over course of the study. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

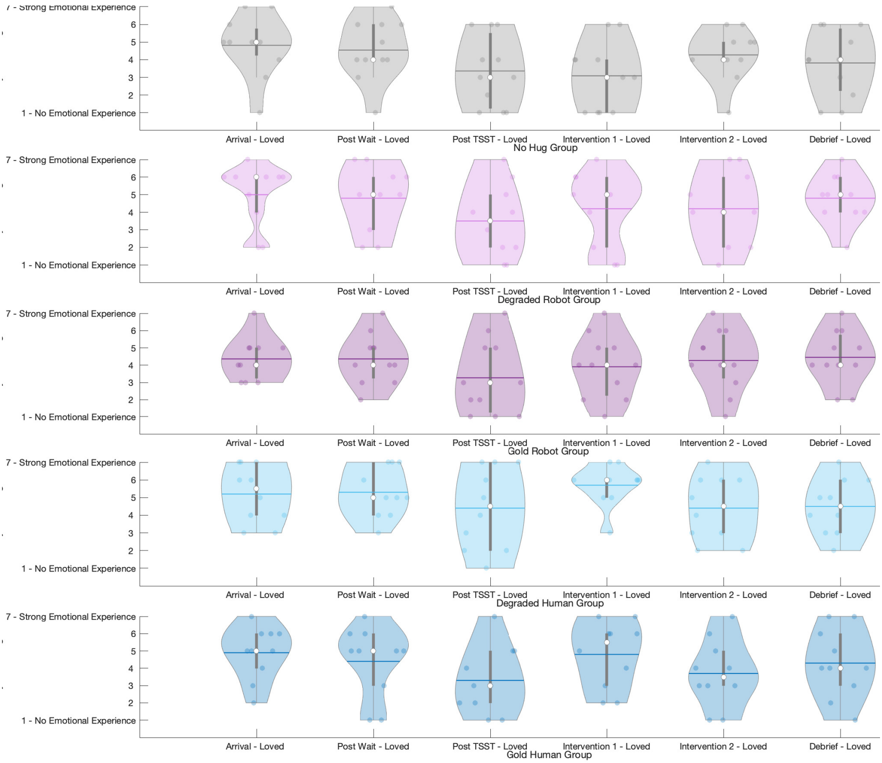


FIGURE 7.16: Violin plots of the participants' responses to the QAS question asking users to assess the quality of how "loved" they felt at each of the six time points the data was collected over course of the study. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

QAS - Seen: Figure 7.17 shows the user responses to how "seen" they felt at the six different time points during the study, separated by which hug condition they received during the first intervention. According to Mauchly's test, the assumption of sphericity had been violated ($\epsilon = 0.66297$), so we report Greenhouse-Geisser corrected results. Overall, there was not a significant change in how "seen" users felt over the course of the experiment ($F(5, 235) = 1.2557, p = 0.28396$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found that users in the no hug condition did not experience any significant changes in how seen they felt at any point during the study. The passive robot hug group also did not experience any significant changes in how seen they felt throughout the experiment. People in the robot hug condition and robotic human hug condition did not experience any significant changes in how seen they felt over the experiment duration. Users in the human hug group experienced a significant decrease in how seen they felt from the first intervention to the second ($p = 0.047699$).

QAS - In Sync: Figure 7.18 shows the user responses to how "in sync" they felt at the six different time points across the study separated by which hug condition they received during the first intervention. Running Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon = 0.71137$), so the Greenhouse-Geisser corrected results are reported. Overall, there was a significant change in how in sync users felt throughout the experiment ($F(5, 235) = 8.181, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found no significant changes in how in sync users felt who were in the no hug condition. Neither the passive robot hug nor the robot hug groups experienced a significant change in how in sync they felt throughout the experiment. Users in the robotic human hug condition felt a significant decrease in how in sync they felt during the debrief compared to how they felt during the initial wait ($p = 0.059467$), and conversely felt an increase in how in sync they felt that approached significance during the first hug intervention compared to the immediately following the TSST ($p = 0.007176$). Users in the robotic human hug group also felt a decrease in how in sync they felt after the second hug intervention ($p = 0.001977$) and during the debrief ($p = 0.014098$) compared to how in sync they felt during the first hug intervention. The human hug group experienced an increase in how in sync they felt after the first intervention compared to after the TSST that

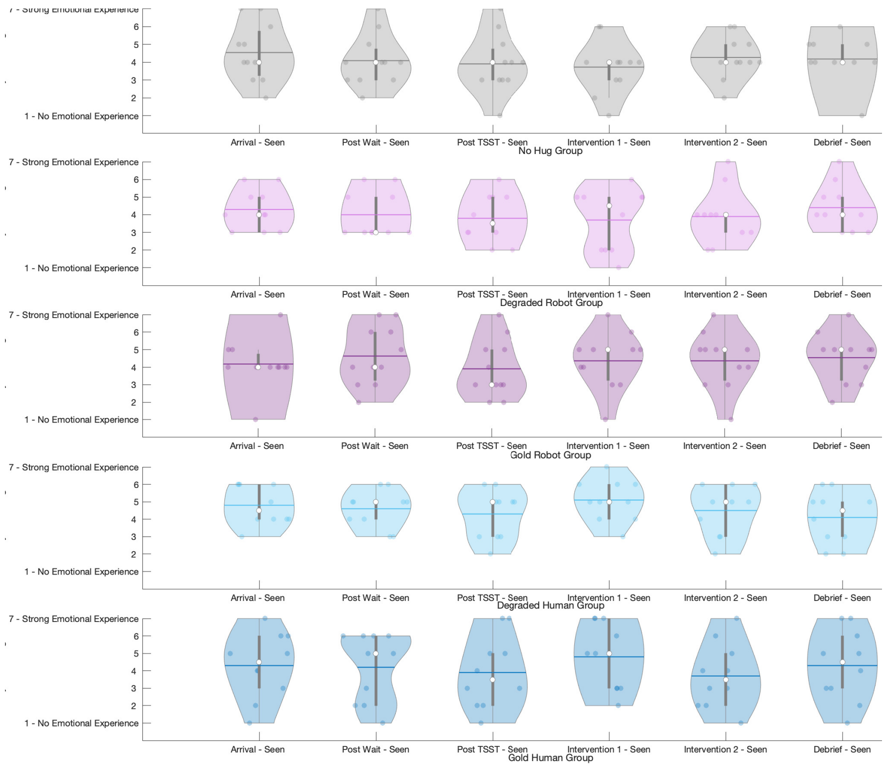


FIGURE 7.17: Violin plots of the participants' responses to the QAS question asking users to assess the quality of how "seen" they felt at each of the six time points the data was collected over course of the study. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

approached significance ($p = 0.080502$). This group also experienced a significant increase in how in sync they felt during the debrief compared to post TSST ($p = 0.016848$).

7.5.2.2 *The Positive and Negative Affect Schedule (PANAS)*

The Positive and Negative Affect Schedule (PANAS) was administered to participants at five different time points during the study: upon arrival, after the TSST, after the first intervention, after the second intervention, and during the debriefing period. Since the PANAS is longer than the QAS, we did not administer it during the post wait as we did with the QAS. Figure 7.19 shows the combined positive and negative self-reported scores by users at each of the five time points during the study, separated by which hug condition they were sorted into for the first intervention method. Again, we used a repeated-measures analysis of variance (`ranova` in MATLAB 2019b) to analyze the differences of the self-reported positive and negative emotions with a significance value of $\alpha = 0.05$.

Positive PANAS: Using Mauchly's test, we found that the assumption of sphericity had been violated for the positive emotions ($\epsilon = 0.85972$), so we report the Greenhouse-Geisser corrected results. Overall, there was a significant change in the positive emotions users felt throughout the experiment ($F(4, 188) = 10.422, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found that there were no significant changes in the positive emotions of the no hug group. Users in the passive robot hug condition showed a significant decrease in positive emotions during the second hug intervention period compared to their arrival ($p = 0.046456$). The robot hug group had a significant decrease in positive emotions reported after the first hug intervention compared to their arrival ($p = 0.0022448$), as well as during the debrief period compared to their arrival ($p = 0.0030963$). The robotic human hug group experienced a significant decrease in positive emotions after the second intervention compared to after their arrival ($p = 0.046456$) and a significant decrease in positive emotions during the debrief period compared to their arrival ($p = 0.046456$). Those in the human hug group had a decrease in the positive emotions reported between the post TSST compared to their arrival that approached significance ($p = 0.063758$).

Negative PANAS: Since Mauchly's test for the negative PANAS indicated that the assumption of sphericity had been violated for the negative emotions ($\epsilon = 0.58227$), Greenhouse-Geisser corrected results are reported.

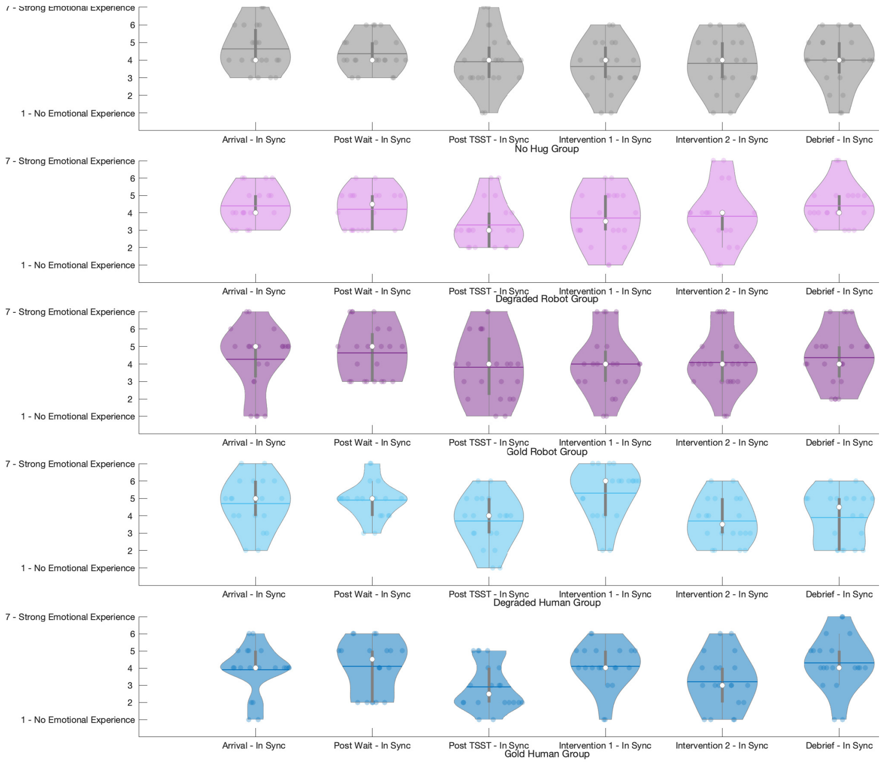


FIGURE 7.18: Violin plots of the participants' responses to the QAS question asking users to assess the quality of how "in sync" they felt at each of the six time points the data was collected over course of the study. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

Overall, there was a significant change in the negative emotions users felt throughout the experiment ($F(4, 188) = 35.232, p < 0.001$). Using a Tukey-Kramer posthoc multiple comparison test taking into account the within-subjects factor of time and the between-subjects factor of hug condition, we found that the no hug group had a significant increase in negative emotions from arrival to after the TSST ($p = 0.0035823$). This group then had significant decreases in negative emotions from after the TSST to after the first intervention ($p < 0.001$), from after the TSST to after the second hug intervention ($p < 0.001$), and finally from after the TSST to the debrief period ($p < 0.001$). The passive robot hug group did not experience any significant changes in self-reported negative emotions throughout the experiment. The robot hug group had a significant decrease in self-reported negative emotions from after the TSST to after their first hug intervention ($p = 0.023704$). The robotic human hug group had a significant increase in negative emotions from their arrival to after the TSST ($p = 0.0038829$). They also experienced significant decreases in negative emotions from after the TSST to after the first intervention ($p < 0.001$), from after the TSST to after the second hug intervention ($p < 0.001$), and finally from after the TSST to the debrief period ($p < 0.001$). Users in the human hug group saw an increase in negative emotions from their arrival to after the TSST that approached significance ($p = 0.051985$). They also experienced significant decreases in negative emotions from after the TSST to after the first intervention ($p < 0.001$), from after the TSST to after the second hug intervention ($p < 0.001$), and finally from after the TSST to the debrief period ($p < 0.001$).

7.5.2.3 *Post-Intervention Ratings*

After each intervention, users were asked to rate the safety, naturalness, enjoyability, social intelligence, and friendliness of the agent they interacted with during the ten-minute session. The no hug group was asked to rate the room they were in based on these qualities. Figure 7.20 shows the user ratings for each of these questions for both interventions they experienced. Since the users answered the same questions at two different time points, we again used a repeated-measures analysis of variance to see how their responses changed from the first intervention to the second.

Safety: The users in the human hug condition rated the first hug interaction (with the human) as significantly safer than the second hug interaction (with the robot) ($p = 0.026497$).

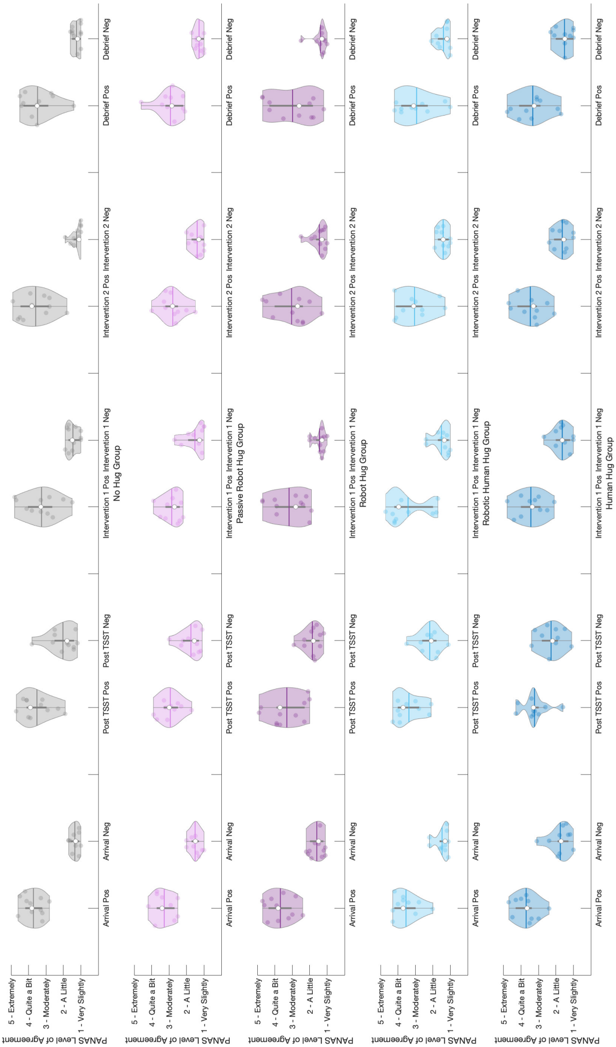


FIGURE 7.19: Violin plots of the participants' responses to the PANAS questionnaire over the course of the study. The ten negative emotions were averaged into a single negative value, and the same was done for the positive emotions. From left to right across the plot time increases. From top to bottom, the hug condition changes from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue).

Naturalness: The people in the robot hug condition rated the first intervention as significantly less natural than the second intervention ($p = 0.033379$), while the users in the human hug condition found the first intervention to be significantly more natural than the second intervention ($p < 0.001$).

Enjoyability: Users in the no hug condition found the first intervention significantly less enjoyable than the second intervention ($p = 0.01416$), while the other groups did not report a significant difference.

Social Intelligence: The no hug group felt that the first intervention was less socially intelligent than the second intervention ($p = 0.025549$), while the other groups did not report a significant difference.

Friendliness: The no hug group ($p = 0.013256$) found the second intervention to be significantly friendlier than the first. In contrast, the human hug group ($p = 0.048573$) found the first intervention more friendly than the second intervention.

7.5.2.4 *Re-Experience Intervention Method*

After users rated the intervention method based on the five adjectives previously described, we also asked the users whether or not they would be willing to experience the intervention method again. In the case of a human or robot agent, we asked if they would be willing to be hugged again by the agent. In the no hug condition, we asked whether or not they would be willing to revisit the room. Figure 7.21 shows the user responses to this question after the first and second intervention periods.

Intervention 1: The top plot in Figure 7.21 shows how users responded when asked whether they would be interested in experiencing the intervention method again. When asked whether they would want to return to the empty room or not, more users (54.5%) in the no hug condition said they would not wish to return. Most users (63.6%) said they would like to hug the robot hugging agent again in the passive robot hug condition. Most users (81.8%) in the robot hug group indicated that they would hug the robot hugging agent again. The same percentage of users in the robotic human hug group said they would be interested in hugging the agent again as in the passive robot hug group, 63.6%. Finally, 80% of the human hug group said they would be interested in receiving another hug from the human hugging agent.

Intervention 2: In the second intervention, all users experienced a robot hug, regardless of the condition they experienced in intervention 1. The second subplot in Figure 7.21, shows how the users responded to being

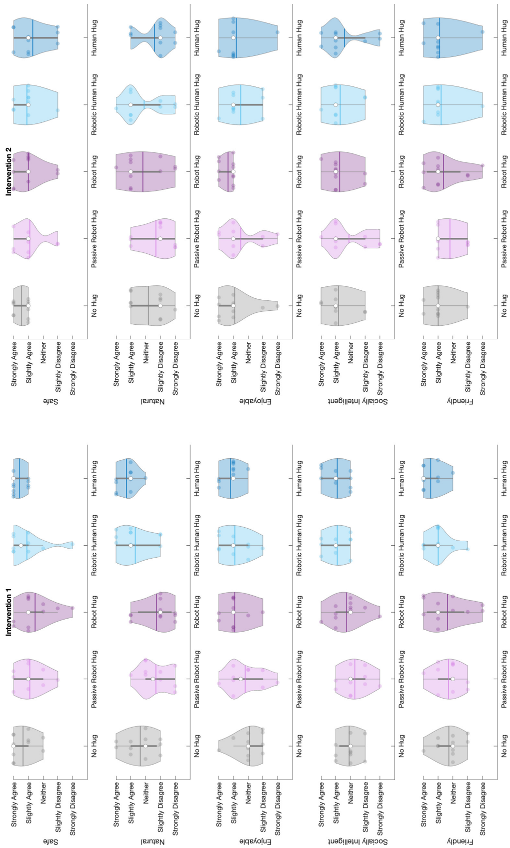


FIGURE 7.20: Violin plots of the participants' responses to the question asking them to rate the safety, naturalness, enjoyability, social intelligence, and friendliness of the agent they interacted with during each intervention. The hug condition are shown in different colors from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue). From top to bottom, the rows show the user responses to "safety," "naturalness," "enjoyability," "social intelligence," and "friendliness." The left side shows the ratings for the first hug intervention, where each group received a different hug condition, and the right side shows the ratings for the second hug intervention where everyone experienced a robot hug, but it is shown based on what hug condition they received in the first intervention.

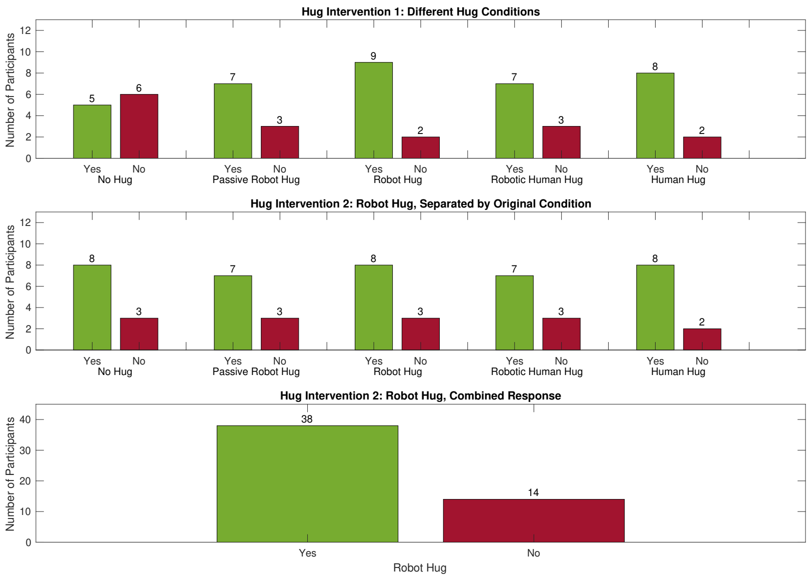


FIGURE 7.21: Bar plots of the participants' responses to the question "would you want to hug this agent agent?". Top: shows user responses after the first intervention method, separated by which intervention they received. Middle: shows user responses after the second intervention method (where everyone received a robot hug), separated by which intervention method they received in the first round. Bottom: shows the combined user responses after the second intervention method (where everyone received a robot hug), not separated by the first intervention they received.

asked if they would like to hug the robot again, separated into what original condition they experienced during intervention 1. Whereas after the first intervention, more users in the no hug condition preferred not to re-experience the intervention method, after the second intervention, 72.7% of users positively responded that they would like to hug the robot again. The same percentage of users (63.6%) who were willing to re-hug the passive robot would be interested in re-hugging the robot. One fewer person would be willing to hug the robot again from the robot hug group after the second intervention method, resulting in 72.7% voting in favor of re-hugging the robot. The same percentage of users (63.6%) who were willing to re-hug the robotic human would be interested in re-hugging the robot. Finally, the same percentage of users (80%) who wanted to hug the human again would also want to hug the robot another time.

Overall, the combined response after the second intervention was 38 users (73.1%) in favor of hugging the robot again, and 14 (26.9%) users were not interested. The total response was close to 3:1, interested in hugging the robot again.

7.5.2.5 *Debrief - Most Beneficial Session*

A final self-reported survey response we asked users was during the debriefing period. We asked users to reflect on both intervention sessions they experienced, and we asked them to state which session was more beneficial after the speech and math tasks. The responses to this question can be seen in Fig. 7.22. More users in the no hug condition, the passive robot hug condition, and the robot hug condition felt that the second session with the robot hug was more beneficial to helping them recover from the stressful speech and math tasks. On the other hand, more users in the robotic human hug and human hug conditions felt that the first intervention was more beneficial to helping them recover. Particularly in the human hug condition, 9 of the 10 users in that group felt the first session was more helpful to their recovery than the second session. However, overall, the number of users who felt the first session was most beneficial compared to those who felt the second session was most beneficial was perfectly split: 26 users preferred the first session, and 26 users preferred the second session.

7.5.3 *Robot Data*

We collected data for each hug users exchanged with the robot to understand the quality of the embrace and how often and which (if any) intra-hug

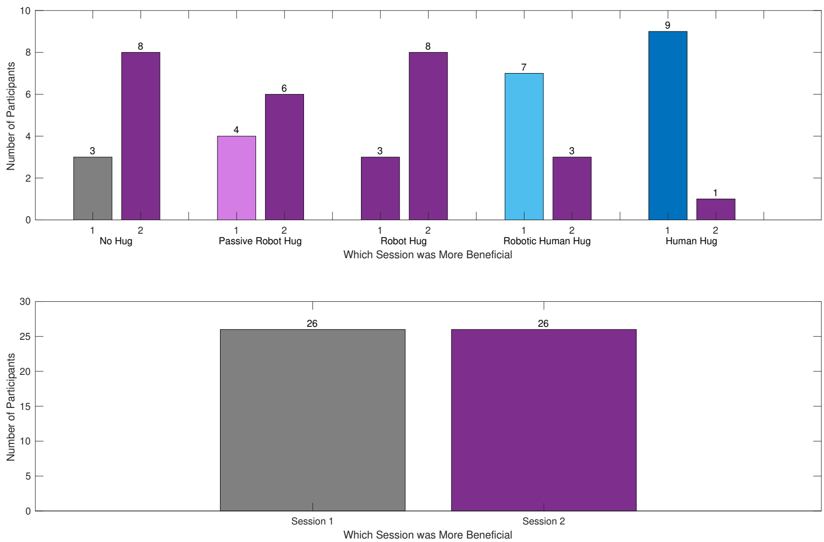


FIGURE 7.22: Bar plots of the participants' responses to the question "Which session was more beneficial for you?". The color of the bars in the top plot represent which hug condition the users received from "no hug" (grey), "passive robot hug" (light purple), "robot hug" (dark purple), "robot human hug" (light blue), to "human hug" (dark blue). The second bar for each group is always dark purple because in the second intervention method all users received a robot hug.

gestures (rubs, pats, and squeezes) users performed on the robot. In Figures 7.23 – 7.27, we show annotated examples of both kinds of robot hugs (passive and active), as well as both kinds of hug durations (timed and pressure).

For each hug, we collected joint angles and joint torques for both the left and right arms. The top two subplots in Figures 7.23 – 7.27 show the data for joints 1, 2, and 3, which correspond to the shoulder lift, shoulder pan, and elbow flex joints, respectively. We also collected and plotted the microphone and pressure signals, which can be found in the bottom two subplots of Figures 7.23 – 7.27. In the microphone signal, while the robot can hear the user making contact with its inflatable chamber, the robot also hears itself invite the user for a hug and the sound of its own arms moving (closing and opening).

Figure 7.23 shows an example of a robot hug with a timed release. In this case, the user did not apply sufficient pressure for the robot to understand that a user was hugging it. Therefore, it used a timed release where it waits three seconds after it finishes fully closing its arms to squeeze and then release the user; as mentioned before, this duration was based on research on human-human hug durations (Nagy, 2011).

Figure 7.24 shows an example of a robot hug with a user-initiated pressure release. In this case, when the user put his arms around the robot, he applied a sufficient amount of pressure for the robot to detect the user's presence. This increase can be seen around 15 seconds on the bottom subplot. The user then held on for a very long time rubbing the robot's back, and the robot responded with many squeezes. The user finally released the robot with a pressure decrease that occurs around 85 seconds.

Compared to the previous two plots, the joint angles and torques in Figures 7.25 – 7.27 do not change as much because the arms do not make contact with the users' bodies. In Figure 7.25, the robot hugs in the passive condition, and the user does not apply enough pressure for the robot to recognize her presence, so the robot releases the user three seconds after it finishes closing its arms.

Figure 7.26 shows the user hugging the robot in the passive robot hug condition. In this hug, the robot invites the user for a hug around 20 seconds, as seen in the microphone signal. The user waits for a while before beginning to walk towards the robot for a hug. The user applies sufficient pressure around 35 seconds, and the robot detects the user-initiated a pressure release hug. The robot then waits until the user releases the pressure around 38 seconds to open its arms and end the hug.

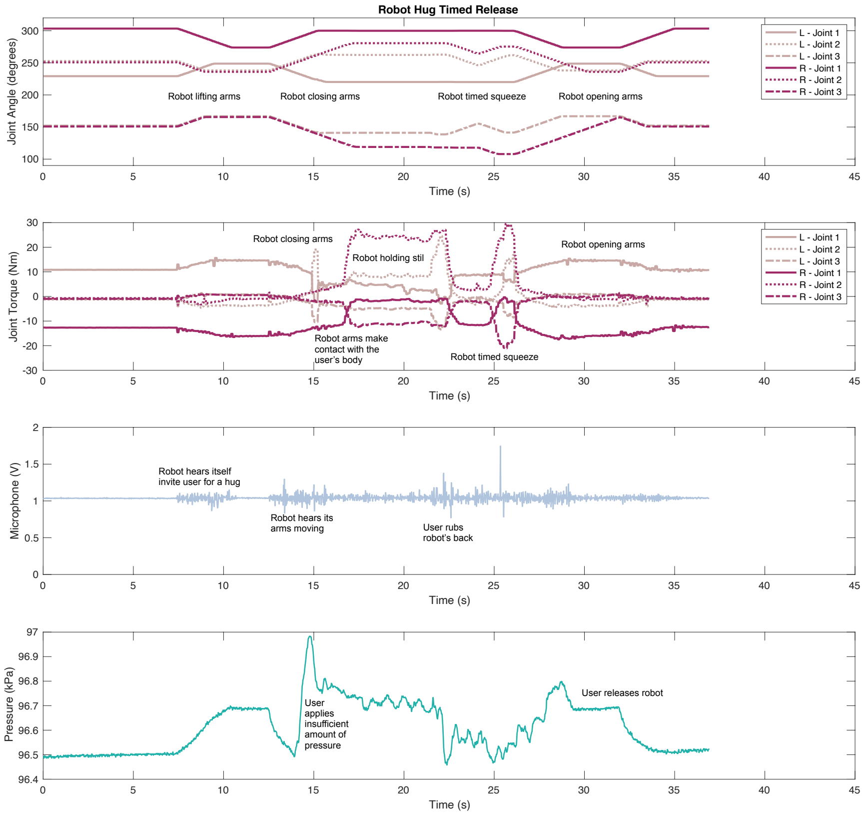


FIGURE 7.23: Example of robot data from a robot in the robot hug condition where the arms adjusted to the user. The user did not apply enough pressure for the robot to realize the user was hugging it, and thus it released the user on a timed schedule. From top to bottom you see: joint angles (joints 1-3), joint torques (joints 1-3), microphone, and pressure signals.

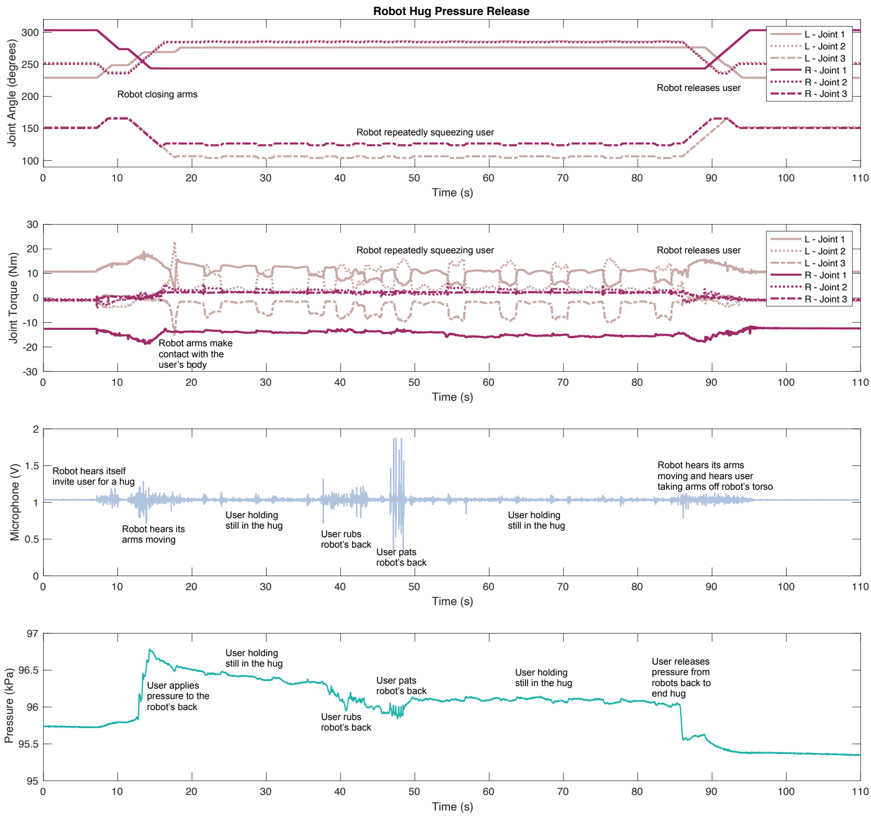


Figure 7.24: Example of robot data from a robot in the robot hug condition where the arms adjusted to the user and it responded to the user's indication of release. From top to bottom you see: joint angles (joints 1-3), joint torques (joints 1-3), microphone, and pressure signals.

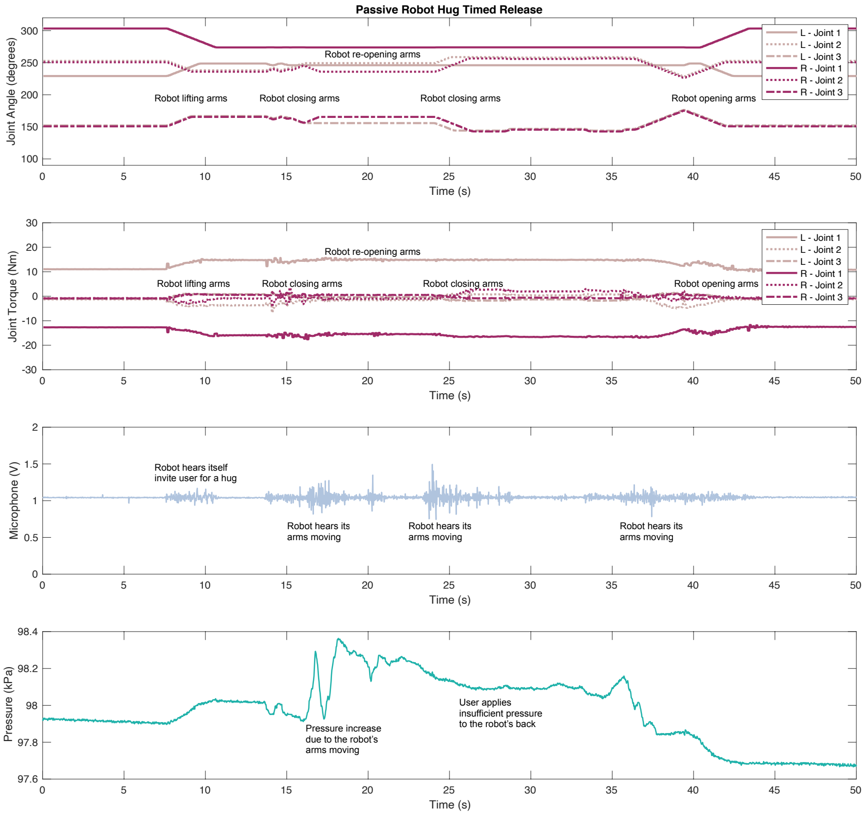


FIGURE 7.25: Example of robot data from a robot in the passive robot hug condition where the arms did not adjust to the user. The user did not apply enough pressure for the robot to realize the user was hugging it, and thus it released the user on a timed schedule. From top to bottom you see: joint angles (joints 1-3), joint torques (joints 1-3), microphone, and pressure signals.

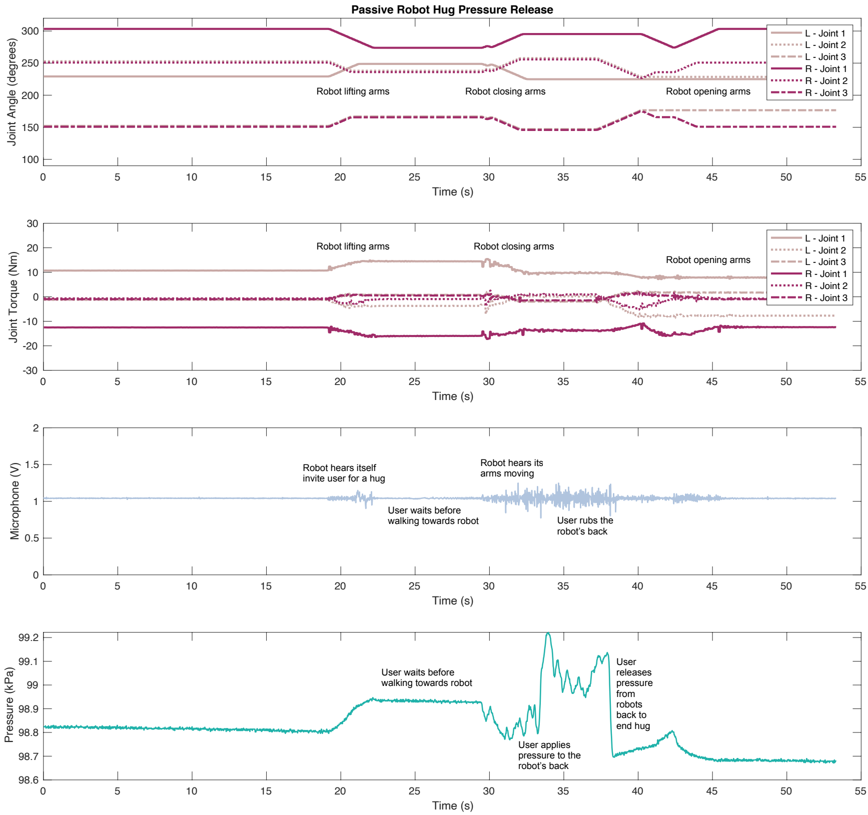


FIGURE 7.26: Example of robot data from a robot in the passive robot hug condition where the arms did not adjust to the user and it responded to the user's indication of release. From top to bottom you see: joint angles (joints 1-3), joint torques (joints 1-3), microphone, and pressure signals.

Figure 7.27 shows a user hugging the robot in the passive robot hug condition. In this case, the user did apply enough pressure for the robot to recognize. However, because the arms did not come into contact with the user's body, he was surprised and released his hands from the robot's back to check if the robot's arms were still moving towards him. When he saw they were stationary, he returned to the hug. Unfortunately, this removal of pressure from the robot's back caused the robot to release the user. This kind of accidental pressure release was most common during the passive robot hug because users were often surprised that the robot's arms did not contact their body.

7.5.4 *Physiological Measures*

During the 3.5-hour experiment, we collected several physiological measures. We continuously monitored participant heart rate and collected two saliva samples at seven different time points. Each time point included one saliva sample to analyze for cortisol and one saliva sample to analyze for oxytocin. We also measured the time it took each participant to provide the same amount of saliva at each of the different time points.

7.5.4.1 *Heart Rate*

As cortisol can have a delayed reaction time of up to twenty minutes depending on each person, collecting continuous participant heart rate gives us a real-time indication of their stress level. It also showed how effective the Trier Social Stress Test was and even allowed us to see which part of the stress test was more difficult for each user. Finally, we used it to see how well our users recovered after the stress test.

While a more sophisticated analysis was not a part of this thesis, the experimenter did investigate user self-reported stress during the TSST and heart rate increase during that time. Anecdotally, she found that men tended to report that they were not very affected by the TSST, when in fact their heart rates told a different story. We plan to investigate this trend more thoroughly to see if it holds true across the male participants in general.

Figure 7.28, shows the heart rate from Participant 26, a male user who had no hug in the first intervention session. Like all participants, he experienced the robot hug in the second intervention session.

Figure 7.29, shows the heart rate from Participant 7, a female user who had a passive robot hug in the first intervention session and then the robot hug in the second intervention session. She was equally stressed by the

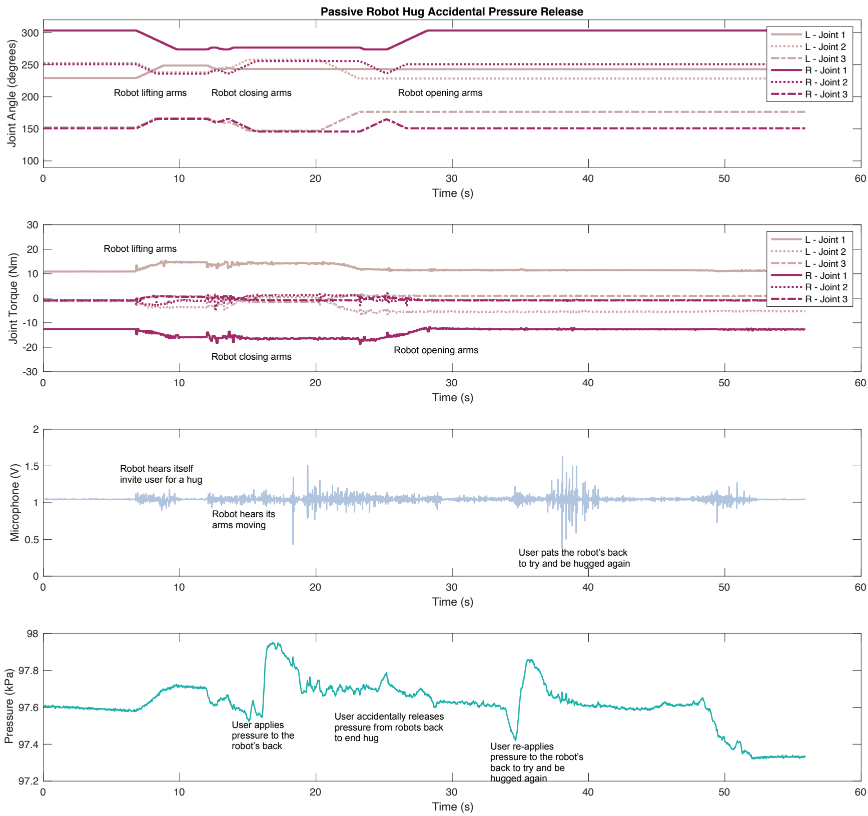


FIGURE 7.27: Example of robot data from a robot in the passive robot hug condition where the arms did not adjust to the user and it responded to the user's indication of release, although the user accidentally release pressure and wanted to continue hugging. From top to bottom you see: joint angles (joints 1-3), joint torques (joints 1-3), microphone, and pressure signals.

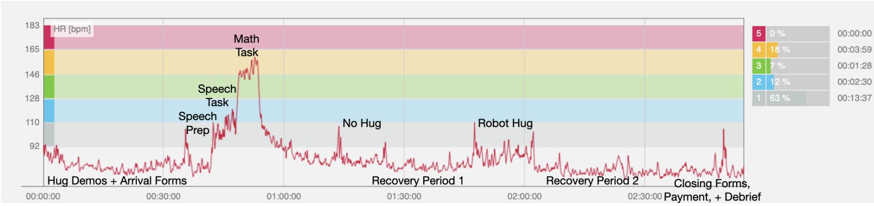


Figure 7.28: Annotated heart rate data from Participant 26 in the no hug condition throughout the entire experiment. Key elements of the experiment timeline are written in text above or below the heart rate.

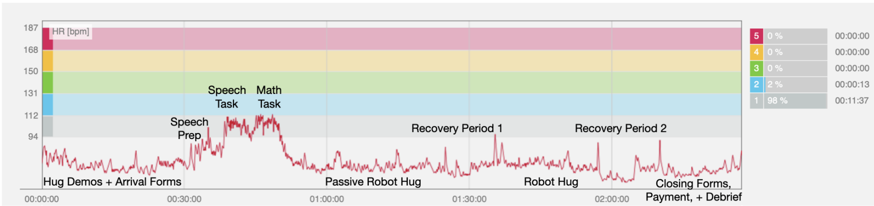


Figure 7.29: Annotated heart rate data from Participant 7 in the passive robot hug condition throughout the entire experiment. Key elements of the experiment timeline are written in text above or below the heart rate.

speech and math tasks. There is a slight dip in heart rate between the two sessions while she received the instructions for the math task. It is also important to note that it is natural for most users to have a slightly higher heart rate during the intervention method than during the recovery methods. During the interventions, the user stands, walks, and interacts with the hugging agent, while during the recovery periods, the user sits still and rests. Heart rate increases with movement, so we expect to see a higher heart rate during the intervention methods than during the recovery periods.

Figure 7.30, shows the heart rate from Participant 22, a male user who had a robotic human hug in the first intervention session, and then the robot hug in the second intervention session.

7.5.4.2 *Spitting Time*

We collected a new physiological measure in this study: the time it takes a user to provide the saliva sample. We believe saliva production is associated

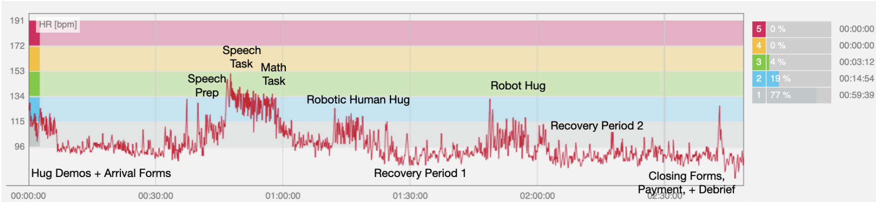


FIGURE 7.30: Annotated heart rate data from Participant 22 in the robotic human hug condition throughout the entire experiment. Key elements of the experiment timeline are written in text above or below the heart rate.

with stress and, therefore, we may see similar trends in the cortisol measurements. The experimenter took note of the time the participant started and stopped providing each of the seven saliva samples.

Differences Between Groups: Using a one-way analysis of variance (`anova1`), we tested to see if there were significant difference between different groups at different time points during the study. Neither at arrival ($F(4, 47) = 1.06, p = 0.3856$) nor after the initial waiting period ($F(4, 47) = 0.49, p = 0.7418$) were there any significant differences between any of the five groups. There were also no significant differences between groups after the TSST ($F(4, 47) = 0.56, p = 0.6932$), the last time point where all groups had the same experience. No significant differences between the groups were observed after the first intervention ($F(4, 46) = 0.59, p = 0.6695$), nor after the first recovery period ($F(4, 45) = 0.83, p = 0.5139$). Additionally, after the second intervention period, there were no significant differences between the hug groups ($F(4, 47) = 0.49, p = 0.7452$), and finally there were no significant differences between the groups after the second recovery period ($F(4, 45) = 0.96, p = 0.4411$).

Figure 7.31 shows the distribution of the participant spit times at each data collection point. Figure 7.32 shows the same data, but a logarithmic transform has been applied to more clearly illustrate the differences in the means. Since the data were not normally distributed, we used the log of the spit time to calculate statistical significance. We again used a repeated-measures analysis of variance with a significance value of $\alpha = 0.05$. Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon = 0.55475$), so we report Greenhouse-Geisser corrected results. Overall there was a significant change in the log of the time it took to spit over the seven different time points in the experiment

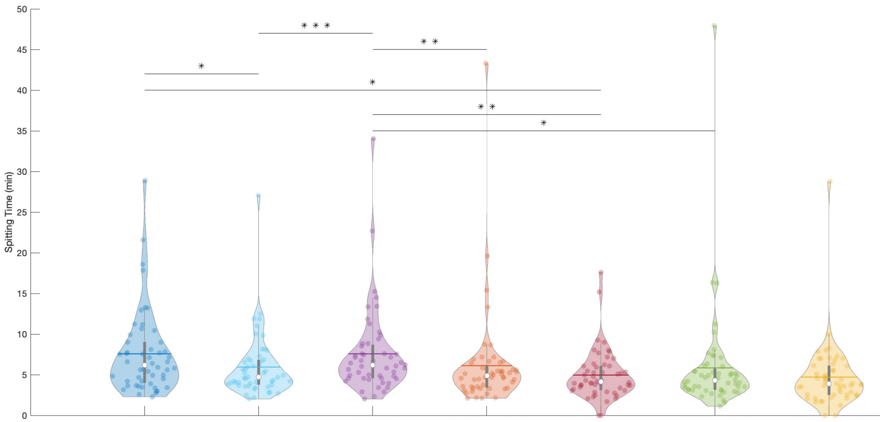


FIGURE 7.31: Violin plots showing the distribution of our participants' spitting time.

($F(6, 252) = 4.5143$, $p = 0.0034372$). In the repeated measures model, we included the within-subjects factor of time and the between-subjects factor of hug condition and Body Mass Index (BMI). We again used a Tukey posthoc multiple comparison test to understand where the significant differences were. Fig 7.32 shows the log of spit times at the seven different time points, but it is not separated by the hug condition. The decrease in the log of time to spit approached significance between arrival and post wait ($p = 0.076701$) and between arrival and post-recovery session 1 ($p = 0.095287$). A significant decrease in the log of spit time was observed between arrival time and post-intervention 1 ($p = 0.020292$). There was also a significant decrease in the log of spitting time after the first intervention compared to after the TSST ($p = 0.015336$), and the decrease after recovery period 1 compared to after the TSST approached significance ($p = 0.075329$).

Differences Within Groups: Figure 7.33 shows the time it took to provide each saliva sample at the seven different time points, separated by which intervention method they received in the first session. The no hug group had a significant decrease in the log of time to spit between arrival and post initial waiting period ($p = 0.025215$), as well as between arrival and post-intervention 1 ($p = 0.0025824$), arrival and post-recovery 1 ($p = 0.04094$), and arrival and post-recovery 2 ($p = 0.020501$). There were no significant differences in the log of spitting time observed for the passive robot hug group or for the robot hug group. The robotic human hug group had

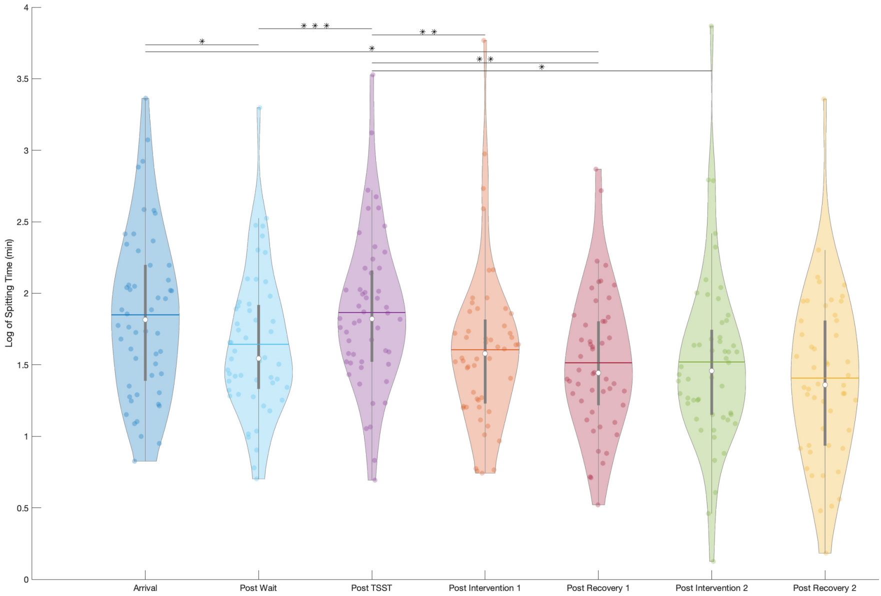


FIGURE 7.32: Violin plots showing the distribution of the logarithmic transform our participants' spitting time.

a decrease in the log of spit time that approached significance between arrival and post wait ($p = 0.09188$), as well as between arrival and post-intervention 2 ($p = 0.075414$). There was a significant decrease observed between arrival and post-intervention 1 ($p = 0.0088237$). There were no significant differences in the log of spitting time observed for the human hug group.

We found that that the interaction between time and Body Mass Index (BMI) had a significant effect of the log of time it took participants to provide saliva ($F(18, 252) = 2.4034, p = 0.01155$). We used four different groups of BMIs, according to the Center for Disease Control and Prevention ("About Adult BMI", 2020), which can be seen in Table 7.5. For *healthy individuals*, there was a significant decrease in log of spitting time between arrival and post wait ($p < 0.001$), as well as between arrival and post-intervention 1 ($p < 0.001$), arrival and post-recovery 1 ($p < 0.001$), arrival and post-intervention 2 ($p < 0.001$), and arrival and post-recovery 2 ($p < 0.001$). There was a significant increase in log of spitting time for healthy

BMI Categories	
BMI <18.5	Underweight
18.5 <BMI <25	Healthy
25 <BMI <30	Overweight
BMI >30	Obese

TABLE 7.5: BMI categories used for repeated measures analysis of variance.

individuals between post wait and after the TSST ($p = 0.0032656$), and a significant decrease between post wait and post-recovery 2 ($p = 0.003524$). Healthy individuals also saw a significant decrease after the TSST compared to intervention 1 ($p < 0.001$), recovery 1 ($p < 0.001$), intervention 2 ($p < 0.001$), and recovery 2 ($p < 0.001$). There was also a significant decrease in log of spitting time between post-intervention 1 and post-recovery 2 ($p = 0.0059383$), as well as between post-recovery 1 and post-recovery 2 ($p = 0.023299$). There were no significant differences in the log of spitting time observed for underweight, overweight, or obsese individuals, possibly partially because these three groups contain fewer users than the healthy group.

7.5.4.3 *Cortisol*

Differences Between Groups: Using a one-way analysis of variance (`anova1`), we tested to see if there were significant difference between different groups at different time points during the study. Neither at arrival ($F(4, 47) = 0.087, p = 0.4911$) nor after the initial waiting period ($F(4, 47) = 1.41, p = 0.2465$) were there any significant differences between any of the five groups. There were also no significant differences between groups after the TSST ($F(4, 47) = 0.64, p = 0.6387$), the last time point where all groups had the same experience. No significant differences between the groups were observed after the first intervention ($F(4, 46) = 0.47, p = 0.7579$), nor after the first recovery period ($F(4, 45) = 0.58, p = 0.6787$). Additionally, after the second intervention period, there were no significant differences between the hug groups ($F(4, 47) = 0.76, p = 0.5594$), and finally there were no significant differences between the groups after the second recovery period ($F(4, 45) = 0.8, p = 0.5302$).

Differences Within Groups: We then used a repeated-measures analysis of variance with a significance value of $\alpha = 0.05$. Since the cortisol data were

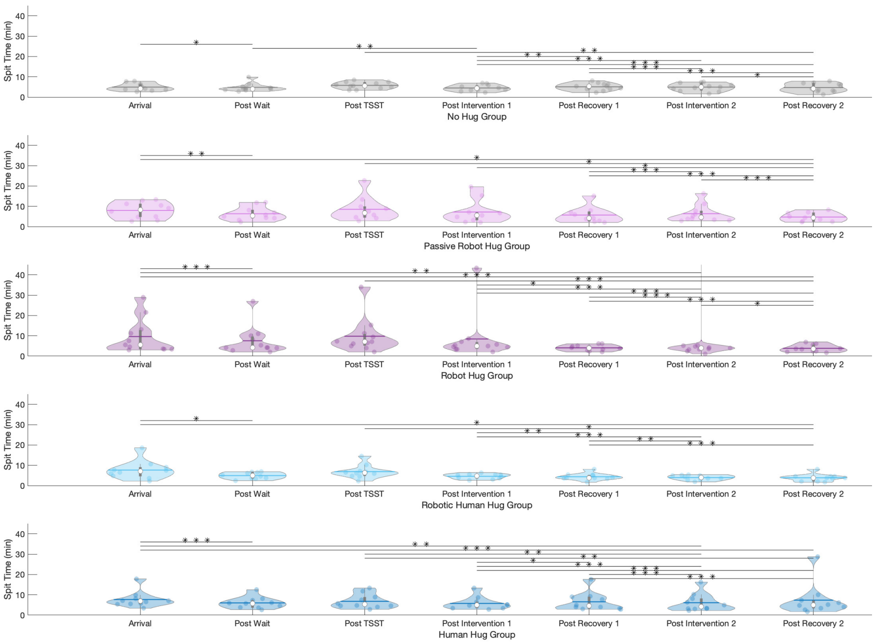


FIGURE 7.33: Violin plots showing the distribution of our participants' spitting time separated by which hug condition they were in.

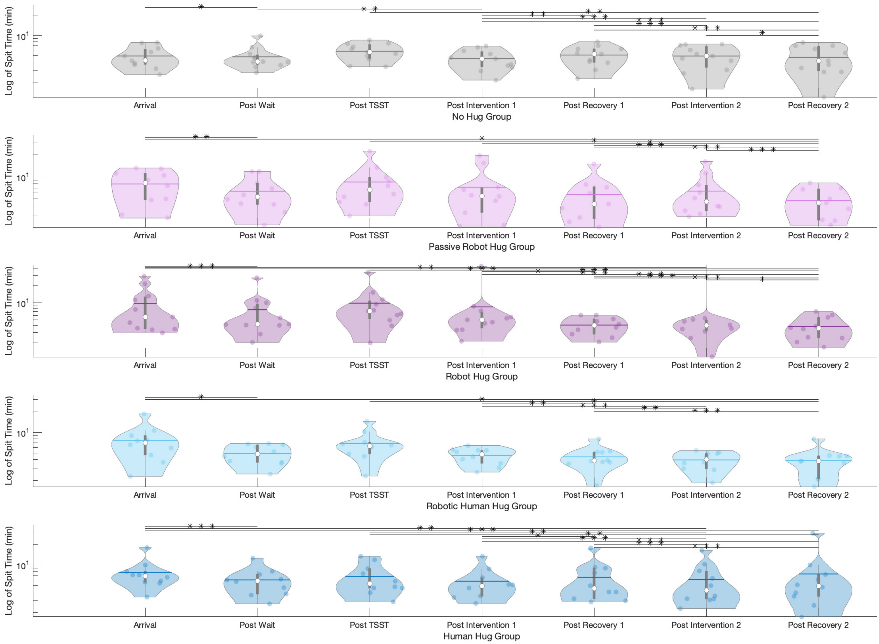


FIGURE 7.34: Violin plots showing the distribution of the log our participants' spitting time separated by which hug condition they were in.

not normally distributed, we used the log of the concentration of cortisol to do our analyses. Mauchly's test once again indicated that the assumption of sphericity had been violated ($\epsilon = 0.3137$), so we report Greenhouse-Geisser corrected results. Overall there was a significant change in the log of the concentration of cortisol in saliva over the seven different time points in the experiment ($F(6, 264) = 30.911, p < 0.001$). In the repeated measures model, we included the within-subjects factor of time and the between-subjects factor of hug condition. Gender, BMI, age, and time of day did not appear to be significant co-factors for cortisol. Although overall, there was not a significant change noticed in the log of the concentration of cortisol in saliva over the different time points when looking with respect to which hug condition users were in ($F(24, 264) = 0.69265, p = 0.85771$), significant individual differences can be observed between time points in each hug condition using a Tukey-Kramer posthoc multiple comparison test.

No hug group: The no hug group had a significant decrease in cortisol from arrival to after the initial waiting period ($p = 0.0024224$). There was a significant increase in cortisol from the post wait to after the first intervention ($p = 0.0088702$). The increase in cortisol between post wait and recovery 1 approached significance ($p = 0.063562$). There was a significant decrease in cortisol between post TSST and post-recovery 2 ($p = 0.042061$). Several significant decreases were noticed in relation to post-intervention 1: post-recovery 1 ($p = 0.0027742$), post-intervention 2 ($p < 0.001$), and post-recovery 2 ($p < 0.001$). A significant decrease in cortisol was also noticed between post-recovery 1 and post-intervention 2 ($p < 0.001$), as well as between post-recovery 1 and post-recovery 2 ($p < 0.001$). Finally, there was a significant decrease in cortisol between post-intervention 2 and post-recovery 2 ($p = 0.016508$).

Passive robot hug group: In the passive robot hug group, a significant decrease in cortisol was observed in the post wait ($p = 0.002189$), and the post-recovery 2 ($p0.012189$), both with respect to the arrival cortisol concentration level. There was also a significant decrease in cortisol concentration in post-recovery 2 when compared to post TSST ($p = 0.012831$). The post-intervention 1 cortisol levels were significantly higher than post-recovery 1 ($p = 0.052113$), post-intervention 2 ($p < 0.001$), and post-recovery 2 ($p < 0.001$). post-recovery 1 cortisol levels were also significantly higher than post-intervention 2 ($p < 0.001$), and post-recovery 2 ($p < 0.001$). Finally, there was also a significant decrease in cortisol concentration in post-recovery 2 compared to post-intervention 2 ($p < 0.001$).

Robot hug group: The arrival cortisol concentration levels for this group were significantly higher than during the post wait ($p < 0.001$), post-intervention 2 ($p = 0.0069519$), and post-recovery 2 ($p < 0.001$). The post TSST cortisol concentration levels were significantly higher than the post-recovery 2 levels ($p < 0.001$). There were also significant decreases in cortisol levels from post-intervention 1 to post-recovery 1 ($p = 0.012744$), post-intervention 1 to post-intervention 2 ($p < 0.001$), and post-intervention 1 to post-recovery 2 ($p < 0.001$). post-recovery 1 cortisol levels were significantly higher than both post-intervention 2 ($p < 0.001$) and post-recovery 2 ($p < 0.001$) levels. Finally, post-recovery 2 cortisol levels saw a significant decrease compared to post-intervention 2 ($p = 0.013007$).

Robotic human hug group: There were significant decreases in cortisol for the robotic human hug group in the post wait compared to their arrival ($p = 0.016508$), and post-recovery 2 compared to their arrival ($p = 0.021518$). Cortisol also dropped significantly in the post-recovery 2 measurement compared to the post TSST measurement ($p = 0.017786$). Significant decreases in cortisol were observed in post-recovery 1 compared to post-intervention 1 ($p = 0.010004$), post-intervention 2 compared to post-intervention 1 ($p < 0.001$), and post-recovery 2 compared to post-intervention 1 ($p < 0.001$). There were also significant decreases in cortisol in post-intervention 2 ($p = 0.001349$), and post-recovery 2 ($p < 0.001$), both with respect to post-recovery 1. Finally, the drop in cortisol between post-intervention 2 and post-recovery 2 approached significance ($p = 0.08202$).

Human hug group: Compared to the arrival cortisol concentration, there were significant decreases in cortisol during post wait ($p < 0.001$), post-intervention 2 ($p = 0.029413$), and post-recovery 2 ($p = 0.0048251$). The post TSST cortisol concentrations were signifiantly higher than both the post-intervention 2 ($p = 0.035151$), and the post-recovery 2 ($p = 0.0024806$). post-recovery 1 ($p = 0.051452$), post-intervention 2 ($p < 0.001$), and post-recovery 2 ($p < 0.001$), all had significant cortisol decreases compared to post-intervention 1. Finally, this group also experienced significant decreases in cortisol in post-intervention 2 compared to post-recovery 1 ($p < 0.001$), and post-recovery 2 compared to post-recovery 1 ($p < 0.001$).

7.5.4.4 *Oxytocin*

Differences Between Groups: Using a one-way analysis of variance (`anova1`), we tested to see if there were significant difference between different groups at different time points during the study. At arrival ($F(4, 47) = 4.92, p = 0.0021$), the no hug group and the human hug group oxytocin

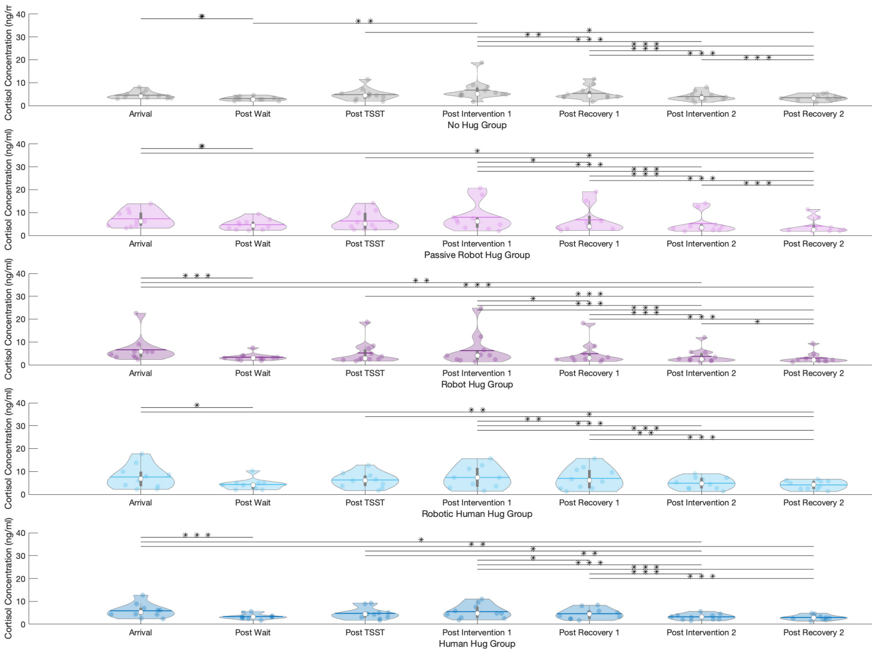


FIGURE 7.35: Violin plots showing the distribution of participants' cortisol concentrations shown at each time point in the study and separated by which hug condition they received in the first intervention.

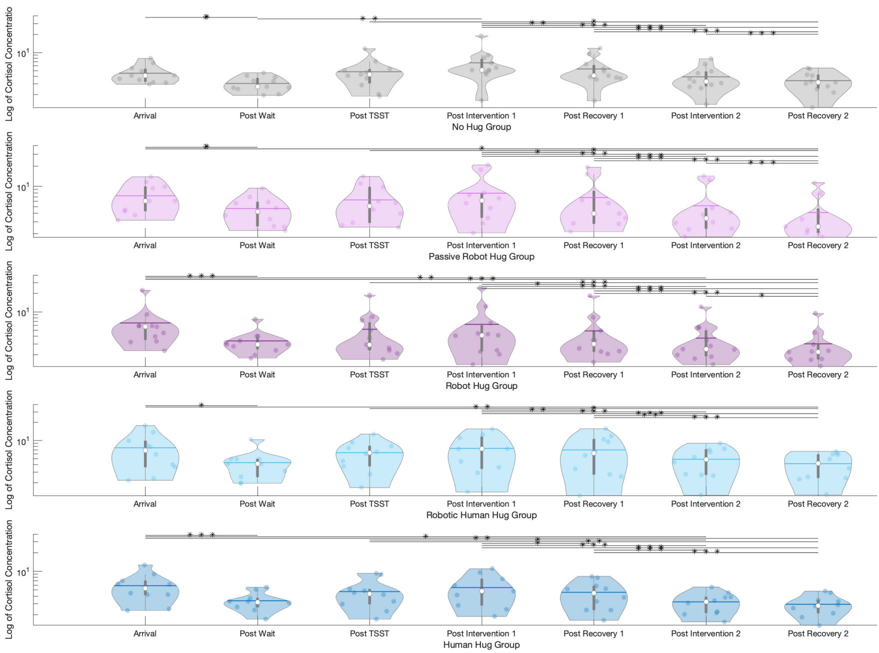


FIGURE 7.36: Violin plots showing the distribution of the log of participants' cortisol concentrations shown at each time point in the study and separated by which hug condition they received in the first intervention.

levels are both significantly lower than the robot hug group. After the wait ($F(4,47) = 3.09, p = 0.024$), the robot hug group had significantly higher oxytocin levels than the no hug group. There were no significant differences between groups after the TSST ($F(4,46) = 1.92, p = 0.1235$), the last time point where all groups had the same experience. No significant differences between the groups were observed after the first intervention ($F(4,47) = 1.71, p = 0.1629$), nor after the first recovery period ($F(4,45) = 1.37, p = 0.2599$). Additionally, after the second intervention period, there were no significant differences between the hug groups ($F(4,47) = 1.62, p = 0.1844$), and finally there were no significant differences between the groups after the second recovery period ($F(4,45) = 1.39, p = 0.2524$).

Differences Within Groups: We then used a repeated-measures analysis of variance with a significance value of $\alpha = 0.05$ to test if there were significant changes across the different time points collected in the study. Since the oxytocin data were not normally distributed, we used the log of the concentration of oxytocin to do our analyses. Mauchly's test indicated that the assumption of sphericity had been violated ($\epsilon = 0.71914$), so we report Greenhouse-Geisser corrected results. Overall, there was no significant change in the log of the concentration of oxytocin in saliva over the seven different time points in the experiment ($F(6,234) = 1.1901, p = 0.31696$). In the repeated measures model, we included the within-subjects factor of time and the between-subjects factors of hug condition, age, and BMI, all of which were significant co-factors. Gender and time of day did not show up as significant co-factors for oxytocin.

Figure 7.37 shows the concentration of oxytocin at all seven different time points collected, separated by which hug condition users were in for the first intervention method.

Hug Condition and Time: In general, the change in the log of oxytocin concentration between hug conditions and different time points approached significance, ($F(24,234) = 1.5011, p = 0.098123$), and using a Tukey-Kramer posthoc multiple comparison test, we see significant differences between different time points. The no hug condition did not have any significant changes in the log of oxytocin concentration levels throughout the experiment. Nor did the passive robot hug condition have any significant changes in the log of oxytocin concentration levels over the course of the experiment. The robot hug condition had an increase in the log of oxytocin concentration levels that approached significance between post-intervention 1 and post-recovery 2 ($p = 0.077251$). The robotic human hug condition did not have any significant changes in the log of oxytocin concentration levels throughout the experiment. The human hug condition had a significant increase in the log of oxytocin concentration levels from their arrival to the post-recovery 2 ($p = 0.049193$). An increase in the log of oxytocin concentration levels approached significance from post-intervention 2 to post-recovery 2 ($p = 0.076142$).

BMI and Time: In general, the change in oxytocin concentration between BMI and different time points approached significance,($F(18, 234) = 1.6838, p = 0.068637$), but using a Tukey-Kramer posthoc multiple comparison test, we see significant differences between different time points. *Underweight* users experienced a significant increase in the log of their oxytocin concentration levels from their arrival to their post-intervention 2 ($p = 0.043193$), as well as a significant decrease from the post-intervention 2 to post-recovery 2 ($p = 0.033318$). *Healthy* users did not experience any significant changes in the log of their oxytocin concentration levels for the duration of the experiment. Neither *overweight* users nor *obese* users experienced any significant changes in the log of their oxytocin concentration levels throughout the experiment.

Age and Time: In general, there is a significant change in oxytocin concentration between age and different time points ($F(12, 234) = 2.1919, p = 0.026778$). Using a Tukey-Kramer posthoc multiple comparison test, we see the significant differences between different time points. The age categories we used can be found in Table 7.6. Two increases in the log of the concentration of oxytocin from the arrival level for the group in *early adulthood* approached significance: post-intervention 1 ($p = 0.065028$), and post-intervention 2 ($p = 0.093506$). There was a significant increase in the log of the concentration of oxytocin from the arrival level to the post-recovery 1 ($p = 0.04216$). The *middle adulthood* group had a significant increase in the log of oxytocin levels from the arrival to the post-intervention 2 ($p = 0.013472$). There was also a decrease in the log of oxytocin levels from post-intervention 2 to post-recovery 2 that approach significance for this group ($p = 0.065655$). There were no significant changes in the log of the oxytocin concentration for the *late adulthood* group at any of the time points measured.

7.5.4.5 *Spitting Time, Cortisol, and Oxytocin Interaction*

Upon preliminary investigation, some users had a significantly high correlation between spitting time and cortisol. This interaction is one that researchers have not yet looked into but makes sense. As the Sympathetic

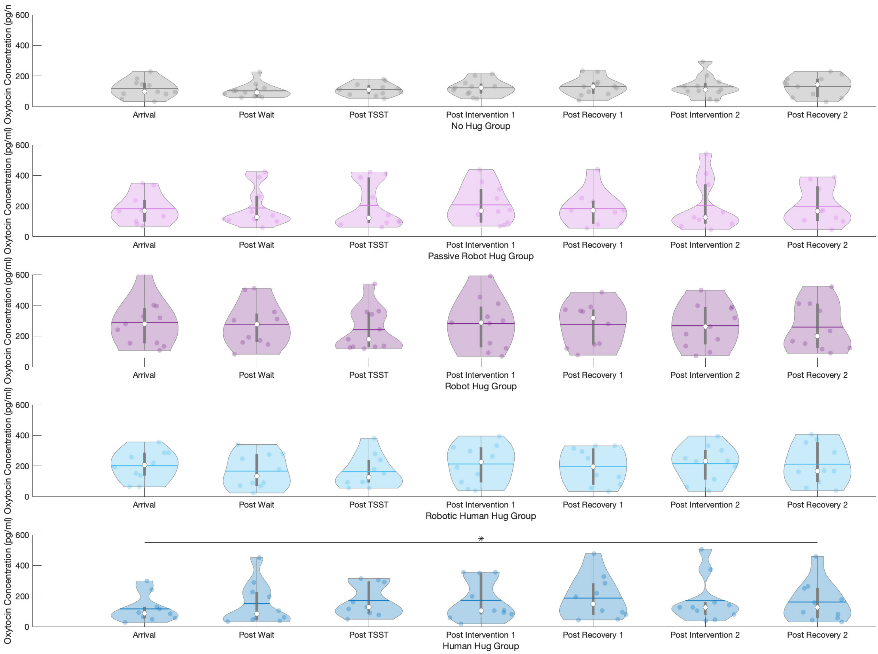


FIGURE 7.37: Violin plots showing the distribution of participants' oxytocin concentrations shown at each time point in the study and separated by which hug condition they received in the first intervention.

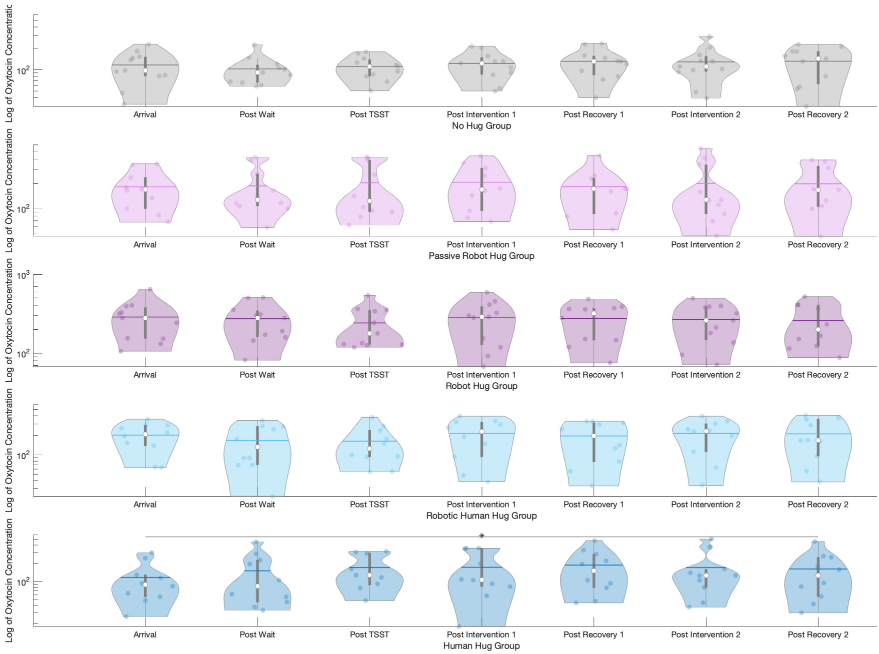


Figure 7.38: Violin plots showing the distribution of the log of participants' oxytocin concentrations shown at each time point in the study and separated by which hug condition they received in the first intervention.

Age Categories	
18 <Age <28	Early Adulthood
29 <Age <39	Middle Adulthood
Age = 40+	Late Adulthood

TABLE 7.6: Age categories used for repeated measures analysis of variance.

nervous system controls the body's "fight or flight" response, it heightens body functions vital to survival and lowers body functions that support relaxation and digestion. As saliva production aids in digestion, we would expect to see that as cortisol increases and users are more stressed, it would become more and more challenging to provide the necessary saliva sample.

We are interested in looking at this more closely in the future, and both measures' interactions with oxytocin. However, for the purpose of this thesis, the interactions between these three measures are out of scope.

Figure 7.39 shows the combined spitting time, cortisol, and oxytocin concentrations of all participants, not separated by intervention method. Figure 7.40 shows the log of these three measures to make it easier to see the difference in means. We see promising trends, particularly between the log of spitting time and cortisol, that further merit investigation.

A more thorough analysis must be done to see if this trend holds true across all participants.

7.5.5 *User Comments and Experimenter Observations*

Though the experiment had a plethora of survey responses, robot data, and physiological measures, another valuable form of insight often comes from user comments made both verbally and in writing. Many users provided written insights into their feelings as part of the QAS. Thus, we share the user comments according to the different time points in the study.

7.5.5.1 *Pre-Arrival and Arrival*

Many users responded positively to online messages and posted flyers. In 83% of emails sent to the investigator to schedule an appointment time, users indicated that they felt hug-deprived in some way. Potential participants would say things like, "I was so excited when I saw this advertisement. It has been months since I hugged another person." Similar

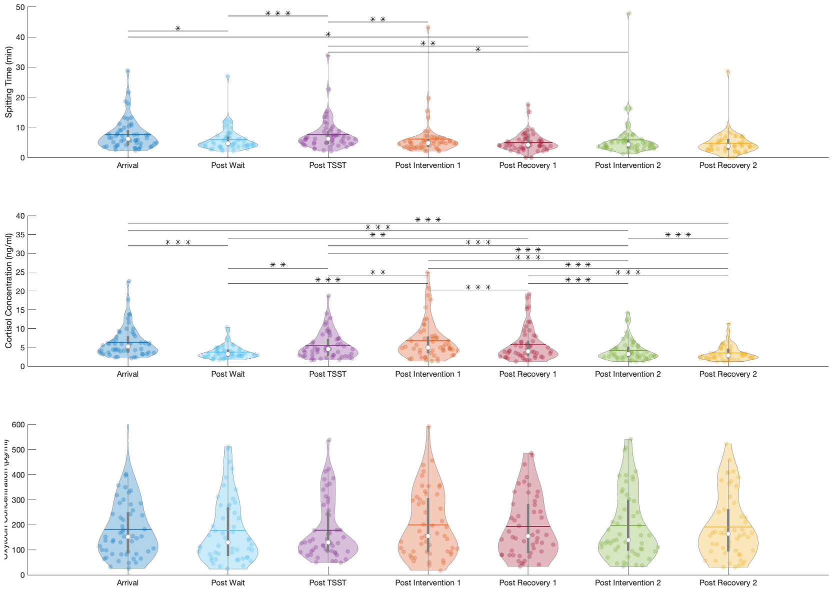


FIGURE 7.39: Violin plots showing the combined cortisol (top) and oxytocin (bottom) concentrations at each time point in the study.

sentiments were often expressed upon arrival in the conference room to begin the experiment.

Two users (3.8%) mentioned they were "scared." One user was scared of the human hugging agent following the hug demonstrations saying, "she did not look human." The second user said they were scared and curious "of the consequences which might occur," possibly referring to the risks section of the informed consent document the experimenter reviewed with users. Twelve users (23.1%) mentioned being very excited about participating in the research study. Only two users (3.8%) mentioned COVID in their arrival surveys. One user mentioned COVID and "the whole place is under lockdown and... there is no social interaction," as a reason for feeling somewhat negative upon arrival. The other mentioned feeling relieved to have received a negative COVID rapid test upon arrival.

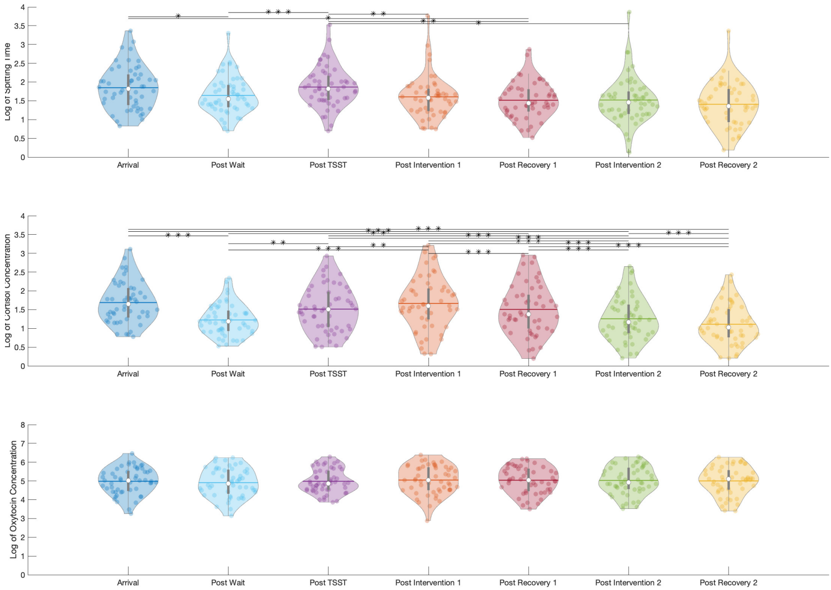


FIGURE 7.40: Violin plots showing the log of the combined cortisol (top) and oxytocin (bottom) concentrations at each time point in the study.

7.5.5.2 *Post Wait*

After the short waiting period, five users (9.6%) mentioned feeling "relaxed," while two users (3.8%) felt "tired." While twelve users were excited for the experiment upon arrival, only two users (3.8%) mentioned being excited for the experiment after the waiting period. Five users (9.6%) felt comfortable.

7.5.5.3 *Post TSST*

The Trier Social Stress Test is designed to induce physiological stress responses in adult participants reliably. The combination of public speaking and a surprise mental arithmetic test in front of an unfriendly panel is bothersome to most people. Therefore, it was unsurprising that the users were the least kind to the experimenter immediately following the TSST. Several users were visibly in tears, many were angry, and some even took their frustration out on the experimenter when she re-entered the room saying things to her like, "that was stupid," or, "I hated that," or even, "why did you make me do that?"

In the written explanation for their negative emotions after the TSST, nine users (17.3%) mentioned feeling very nervous or anxious during the speech and math tasks. Seven users (13.4%) specifically mentioned the surprise element or lack of preparation as to why they felt negative at the time. Twelve users (23.1%) referenced perceived poor performance as the reason for their negative feelings, saying things like, "I am not good in math," or "I did not perform well." Ten users (19.2%) specifically mentioned that they found the task stressful.

In addition to rating their emotional status immediately following the TSST, we also asked users to rate how well they thought they performed in each section of the TSST (speech and math tasks). We also had all three members of the panel of experts rate users on how well they performed each individual task. For this thesis, we did not systematically evaluate the self and panel evaluations of the users' performance. The experimenter did look at these evaluations as the study transpired. The experimenter observed that male users tended to overstate their performance compared to the panel evaluation, while women tended to think they performed worse than the panel did.

7.5.5.4 *Post-Intervention 1*

Until this point in the study, all users had the same experience. At Intervention 1, the users were split into five different groups and had different experiences before coming back together during the Recovery 1 session; from there on to the end of the experiment, all users had the same experience.

No Hug: Of the 11 users in the no hug group, 4 (36%) people had positive comments about the empty room using words like "calming," or "relaxing." Two users (18%) did not share any additional words thoughts about the empty room. Two more users (18%) shared neutral feelings about the room using things like "a feeling without words," and "research area with a lot of materials, seems productive." Two users (18%) were focused on the next element of the experiment. Only one user (9%) explicitly mentioned reflecting on the stress test, saying, "I feel better because I have more time to think that even the previous task was not good, I do not have to feel bad with myself."

Most users were confused upon entering the room and seeing that they were completely alone. Six asked if they were allowed to speak with the experimenter during their ten minutes in the room. Users were told that they could ask if they had a specific question regarding the study, but otherwise, they should spend time in the space. After only two to three of the ten minutes, most users (7 of 10) sat on the chair in the corner of the room, indicating they were finished exploring the space. Five users asked the experimenter, "so... this is it?" When explaining why they would or would not want to revisit the empty room, the main reason for saying no (6 of 11 users, 54% of total group) was because users felt it was boring.

Passive Robot Hug: Two users (20%) mentioned that they felt comforted by the robot immediately after the stressful experience of the TSST. They felt that the robot helped them feel better. One user (P3) even said, "I feel better. I almost felt like a real hug, specially when I closed my eyes." Four users (40%) of users positively mentioned the robot's warmth, while one user (10%) felt the robot was too hot for her liking. Two users (20%) mentioned that they did not feel the robot was very natural, and two other users (20%) mentioned that they did not feel the robot was a very good hugger in this condition, with one user (P17) elaborating, "Not a great hugger. No finesse, no style, no ability to reflect any aspect of the hug (other than length....sometimes)."

When explaining why they would or would not want to hug the passive robot again, most people thought it was an enjoyable experience 5 (50%), and two hypothesizing that perhaps the robot could hug better the next time (20%). In explaining why he would want to receive another hug from the robot, one user (P7) said, "I feel comfortable. He was warm and in some point I just want someone to hug me to reduce my stress. I want to feel the support of someone else."

Robot Hug: In the robot hug condition, the robot fully embraced the users in a custom hug for their body shape. As they did with the passive robot, several users (5 of 11 - 45.45%) commented positively on the robot's warmth, with one user in particular saying, "It has warmth in it. For me, it felt like I was hugging my mother as per se." Three users (27.27%) commented positively on the robot squeezing them, saying, "It felt nice. The hug was warm, and the agent even hugged me tighter sometimes. It felt natural" (P33). Three users (27.27%) specifically mentioned they liked the ease of hugging the robot and appreciated "that the robot was attentive to small cues like pushing back to end a hug or releasing the arms" (P53). When reflecting on the overall ten-minute interaction with the robot, one user (P57) said, "The hugging made me become really quiet and peaceful."

When explaining why they would or would not want to hug the robot again, most users mentioned the excellent quality of the embrace and the warmth combining to create a pleasant and inviting experience (8 of 11 - 72.7%). Participant 2 even said, "It's comfortable experience, the robot intelligence is good to know when you are done with the hug."

Robotic Human Hug: In the robotic human hug, the human agent was dressed and behaved the same way as the robot in the robot hug condition. We did our best to make the robotic human hug condition as similar as possible to the robot hug condition, with the only difference being that one was alive and one was a robot. Two users (20%) in the robotic human hug condition remarked upon hearing the agent's heartbeat and breath and that these two elements were helpful. Participant 49 said, "Feeling the heartbeat and hearing the breath of the agent made me feel safe and helped me to calm down." Three users did not feel it was very natural, with one mentioning it was because the actor was a strange person, while the other two did not elaborate why they felt uncomfortable. One user (P51) compared this agent to her mother, saying, "I felt happy when hugging the human agent because I felt like hugging my mom." Participant 12 commented positively on the hugging agent squeezing during the hug, saying, "I like strong and tight Hugs. It was a friendly tight hug."

In explaining why they would or would not want to hug the robotic human agent again, two uses said they would not because they are shy and hugging a stranger makes them uncomfortable, while five users mentioned wanting to hug her again because of the positive emotions they associated with her hug like, "safe and loved" (P49), "feeling of being cared [for] and friendliness" (P50), "feels relaxing" (P51). Participant 12 said, "Hugs give me emotional peace," while Participant 35 said, "Felt like a grandmother, would definitely do it again."

Human Hug: Three users (30%) mentioned they felt hugging the human agent was "comfortable," "peaceful," and "warm." One user (P4) disagreed and felt that "I think that warmth was missing." Five users (50%) mentioned that the act of hugging is calming to them, regardless of the hugging partner saying things like, ""Hugging is so natural and essential for humans" (P9). Participant 28 shared a preference for long hugs saying, "Longer hugs made me feel better compared to shorter ones." Just like the robotic human hug reminded some of their mothers or grandmothers, the human hug condition also reminded users of their family members, with Participant 36 saying, "It was just a feeling of a warm hug like, I used to do with friend or family."

Many users would want to hug the human agent again, saying things like, "it felt comforting" (P28). Five people (50%) would want to hug the agent again because they missed and enjoyed the feeling of being hugged. One user (P9) specifically liked the human agent saying, "She is careful and warm, her eyes are full of love, tenderness. It felt so good to be in her arms and connect."

One user (P55) specifically mentioned the social anxiety that can sometimes be associated with hugs with other people saying, "I find it a little difficult to make eye contact with the agent. And I felt a little uncomfortable because I had the feeling my hugs were too long and that the agent was already tired of hugging me." Participants 4 and 44 also felt that hugging an unknown person was an unpleasant element.

7.5.5.5 *Post-Intervention 2:*

During Intervention 2, all users hugged the robot in the robot hug condition, regardless of their condition in the first intervention.

Post Hug Comments: Some users felt that hugging the robot in this condition was similar to hugging another person. Participant 12, who was in the robotic human hug condition for the first intervention, remarked that this robot hug "wasn't very much different than the human agent." Similarly, Participant 19 said, "I felt the robot is giving a hug like a human. Though it is not completely natural but I enjoyed the hug." Two users (3.8%) mentioned missing the heartbeat, breath, and eye contact from the human hug condition. Three users (5.7%) (including the one who missed the human heartbeat) mentioned that they could hear the sound of the computer in the robot's head and that the noise made the hug seem unnatural.

Several people compared it to a human hug, and particularly mentioned how this robot could be helpful during times where physical distance is required, "It gives you a feeling of a human, which can be helpful in these times (pandemic)" (P36). Participant 48 agreed, saying, "It's good to have being hugged tight again after so long." Many users (20 out of 52 - 38.5%) mentioned overall enjoyment of the entire hugging experience, with Participant 3 saying, "I was lost in positivity while hugging :)," or like Participant 30, who said that hugging the robot, "Felt a sense of relief stress." Lots of users (6 out of 52 - 11.5%) commented positively on the robot squeezing them, saying things like, "I liked when he squeezed a little tighter, and it felt warm and cozy" (P16).

While many users like the squeezes, some users commented that it squeezed them too tight. Participant 47 felt the squeeze made the hug less pleasant for her preference. Three users (5.8%) wanted additional padding added to the arms to make them softer. One user (1.9%) wanted the voice to be less robotic.

Some users were initially nervous or hesitant about hugging the robot, like Participant 31, who received no hug in the first intervention period, and when reflecting on his hug with the robot, said, "Was nervous or skeptical at first, but once I got comfortable, it felt very relieving, felt like I could ease some of my burden with a hug." Another user (P50) mentioned that she was more comfortable hugging the robot than the human, "I can hug for longer because it's a robot, I don't feel shy." Participant 20 specifically mentioned appreciating the responsiveness of the robot, saying, "It was fun, hand movement felt different, which was intriguing and made the experience new, every time I hugged the robot."

Why or why not willing to hug again?: When answering whether they would or would not want to hug the robot again, the main reason users said they would be unwilling to hug the robot again was because they had already hugged it a sufficient number of times and feel they understand the "full capacity of the robot," (P32) and thus do not need to hug it again. Four users (7.6%) wanted to hug the robot again because "it was fun to hug the robot." (P20). Ten people (19.2%) positively mentioned the "warmth," seven (13.5%) called it "friendly" or that it reminded them of their friends and four people (7.6%) thought that it was "relaxing" to hug the robot, and mentioned these as reasons they would want to hug the robot again.

QAS: Most users (40 out of 52 - 76.9%) did not feel any negative emotions after the second intervention. Many of these users mentioned that the experience of hugging made them feel better, with one (P36) saying, "It feels the same as you hug a person. The warmth of the robot makes me feel like I am hugging a person."

Comparatively, only nine users did not feel any positive emotions following intervention 2. One user (P57) even said, "It becomes nicer to hug the robot the more you hug him." Participant 3, like five others (11.5%), was pleasantly surprised by how enjoyable hugging the robot was, saying, "I loved the warm hug. I was positively surprised and flowing. Felt great!" Thirty-six users (61.5%) mentioned some aspect of the second intervention and hugging the robot as the reason for their positive emotions. In this space, seven users (13.5%) compared their first experience with the robot to the second one, saying they felt the second intervention was "Better than the first time" (P40).

Six users (11.5%) also mentioned that they did not notice a difference between the human and robot agent saying and even going so far as to as the robot reminded them of a friend or family member. Participant 12 said, "I didn't find as much difference than human agent." Participant 23, who was in the human hug condition for the first intervention, said hugging the robot "was a positive experience, felt warm and relaxed. I did felt my mum hug here, and it was nice." Four users who had the no hug condition mentioned specific people the hug reminded them of, like, "It feels similar as when I hug someone..." (P14), or, "felt like I was hugging a close friend or a family member. Reminded me also of my grandparents - strong and warm" (P31).

7.5.5.6 *Debrief*

QAS: Forty-eight out of fifty-two users (92.3%) felt no negative emotions by the end of the experiment. Five users (9.6%) felt calm or relaxed, five users (9.6%) mentioned feeling tired from the length of the experiment, and three users (5.8%) mentioned being glad the study is almost over. Four users (7.6%) mentioned that the number of saliva samples was challenging for them. Nine users (17.3%) felt no positive emotions when filling out the QAS during the debrief session. Some users (12 out of 52 - 23.1%) mentioned they were happy to be contributing to science by participating in the experiment. Five users (9.6%) felt positive because they were nearing the end of the experiment.

Why more beneficial: Overall, five people (9.6%) who said they found the first session more beneficial than the second session mentioned that it was because it came immediately after the stressful situation.

No Hug: Of the three people in the no hug condition who stated that they felt the first session was more beneficial than the second, all three indicated that when they are stressed, they prefer to handle it alone. One mentioned lying on the floor; another mentioned not being bothered and given time to relax. The eight others who said the second session was more beneficial than the first mentioned that having something to do and someone/something to share the emotions with was more helpful than being alone. Participant 31 says, "Robot was more interactive, it is better to be with someone and share it out."

Passive Robot Hug: In the passive robot condition, of the four users who found the first session more beneficial than the second, three of those users mentioned that the hugs coming immediately after the stressful experience were what made that session feel more beneficial. One user felt that the robot in the second session hugged her tighter than it did in the first session, and thus she could feel the hardness of the arms, which made her prefer the first session. There were two main reasons why the remaining six users felt the second session was more beneficial than the first. The first reason was that they were more familiarized with the robot (2 users) and thus felt more comfortable around it during their second interaction. The second reason is that they appreciated the arms closing fully around them into a custom embrace, saying it felt more "human" (P54) when it touched them (4 users). As Participant 27 says, "The hugs in the 1st session felt incomplete."

Robot Hug: Of the three users who preferred the first session with the robot hug than the second session, all three indicated the reason for their preference being that the hugs came immediately following the stressful experience. Of the eight users who preferred the second session in the robot hug condition, some users mentioned more "trust" (P57) in the robot the second time, and others mentioned they felt more relaxed, which in turn made the experience more enjoyable (P10). Overall, however, each of the eight users mentioned feeling more comfortable around the robot with increased interactions. Participant 53 for example said, "The first session was understand[ing] how it work[ed] and it was a little unfamiliar. The second session was much warmer and better. It went as I expected it. The robot held me as long as I wanted."

Robotic Human Hug: Seven of the users in the robotic human group found the first intervention session more beneficial than the second. One user did not feel much difference between the two sessions, so he felt the first one was more beneficial because it came after the stress. Four users mentioned the human aspect, that it felt "real" or "human," and that the agent put meaning behind the hug that made the user feel "loved and cared for" (P49). The last two users who preferred the robotic human felt that it reminded them of their family members. Three of the users in the robotic human group preferred to hug the robot. Two users said this was because they felt uncomfortable hugging a stranger and that hugging a robot was also more interesting. The third user said it made her more comfortable and relieved her stress.

Human Hug: Nine out of ten users in the human hug condition preferred the first intervention session over the second. Two of these users mentioned that it was only because the hug came immediately after the stressful experience. Five mentioned that they "feel connected with human touch," (P44) and that they were excited they "got to hug a person after so long" (P4). Two users mentioned the human hugging agent's breath feeling calming, with Participant 55 saying, “I found the voice, the heartbeat, the breathing of the person very soothing.” One of those users said the human agent reminded her of her mother (P23). The one user who preferred hugging the robot said, “I am not comfortable hugging a human stranger again and again. But hugging a robot is quite relaxing and fun” (P45).

Additional Comments about the Robot: Five users requested softer arms in the next version of the robot. Two users wanted the robot’s voice changed to be more human-sounding. Two users wanted more robot movement: one wanted the robot to move to the user for a hug, and another wanted the robot’s arms to move more. Three users thought the robot was too hot. One user wanted a stronger hug, one user wanted a human face instead of an animated face, and two users wanted a more human body shape. One user mentioned the desire to move, remove, and replace her hands on the robot’s back without the robot thinking she wanted to be released from the hug.

Thirteen users mentioned liking the idea of a hugging robot and that it could really help people. Participant 23, in particular, said, “I think it is a very good idea and can be really extended in the future, to make anyone feel safe and cared about.” Two users positively mentioned the robot’s interactivity, explicitly mentioning the squeezes saying things like, “It was really interesting, specially how the robot reacted to rubbing his back, etc. It was a nice tight hug” (P33). Participant 55 addressed feeling more comfortable the more interactions a user has with the robot saying, “I can see people relating more and more to the robot each time. It kind of grows on you.” Participant 51 summed up her feelings simply, “I loved hugging the robot.”

Self-Soothing Strategies Used: Our participants used various self-soothing strategies to calm themselves down after the stressful speech and math tasks. Commonly reported techniques included: breathing deeply, using positive thinking, planning for the future, telling themselves their performance did not matter or distracting themselves. Twelve users mentioned taking deep breaths and focusing on breathing exercises. Seven users used the power of positive thinking to help themselves feel better about their performance. For example, Participant 3 said, “I was thinking about my other strengths, not [my] weakness for math.” Two users said they did not need to do anything to calm themselves down, as they did not find the speech and math tasks stressful (P47 and P57). Four users took their performance as motivation to start planning for improvement in the future (be more prepared, or be more spontaneous). For example, Participant 7 said, “I know that I really need to be more prepared. Furthermore, I thought this kind of experiences are important to improve your skills and go through your years," while Participant 44 said this experience motivated him to "start working hard to achieve my goals in my life." Fifteen found comfort in telling themselves that their performance in the speech and math tasks does not matter, that it was not an actual job interview and that they will not see the panel members again. Finally, eight users said that distracting themselves, so they did not think about their performance, helped them. In particular, seven of the eight mentioned that the hug sessions were helpful to get their minds off the previous activity, with Participant 12 saying, "[the] hugging session helped me relief my nervousness and stress which I gained during speech and math sessions."

How usually handle stress: Because we requested that users lock their cell phones away for the research study duration, when we asked how they usually handle their stress (outside of the study), there was a more comprehensive range of answers. Standard techniques include: talking to a loved one, drinking something, eating something, taking a nap, exercising, hugging, positive thinking, preparing for the future, distracting themselves, having a good cry, or spending time in nature. Twenty-one of fifty-two users (40.4%) of our participants said that when they feel stressed, the best thing to do is talk to loved ones. Two users specifically mentioned their mothers, three mentioned their partner/spouse, and the others mentioned close friends or other family members. Four people (7.7%) mentioned they like to drink something warm, such as tea, hot chocolate, or coffee. Nine users (17.3%) say they like to eat something when they are upset, usually mentioning chocolate or junk food. Six users like to take a nap to try and sleep off their bad feelings. Twelve users (23.1%) said some form of exercise helped get over feeling stressed - yoga, meditation, running, walking, and sports were all mentioned. Nine users (17.3%) said hugging is a typical way for them to alleviate their stress. Seven people use the power of positive thinking to change their mindset and help them calm down. Six users like to go over the situation, figure out what went wrong, and think about fixing it or improving it in the future. Participant 32 says, "I write down the mistakes I have made and think and the ways I could have avoided it and find solutions," while Participant 28 said, "I think about what happened, evaluate if my actions were right or how I'd do it differently next time and try to learn from it." Many users mentioned that having a distraction helps them decompress and destress. People use different forms of distraction: nine users like to listen to music to help them destress, fifteen people find watching a movie or a television show calming and relaxing, one user prefers to read, and one user likes to smoke. Four people mentioned using different deep breathing techniques; three like to have a good cry, and three like to spend time in nature.

Reason for Participating: Everyone brings their own reasons for wanting to participate in a research study. It is important to understand those motivations in case they can affect the study results. Six users were interested in visiting the Max Planck Institute for Intelligent Systems and saw participating in this study as an opportunity to do that. Thirteen users were interested in contributing to new research. Nine users mentioned curiosity as a primary driver for them signing up. Fifteen users specifically mentioned wanting to participate because of the hugs. Three separate users mentioned that they have been feeling a lack of social and physical contact due to the pandemic, and this was a way for them to get some kind of affection and social interaction. One user said, "The fact that I have been longing for a bit of affections gestures motivated me more when I found out this has something to do with hugging," (P48), while another said, "During pandemic is less physical contact between people because we are afraid about the virus propagation. But also mental health is important" (P2). Twelve users participated primarily because of the monetary compensation. Six users were motivated to join the study because they were interested in the prospect of testing out new technology.

Participant 26 had a specific reason for participating. He said, "I read through an article where technology development in self-driving cars were not able to recognize people of my skin colour. Because the algorithms were written without our skill colour being considered. For this reason, I decided if technology is being developed for the future, I could help bring about diversity in its development."

Additional Comments: The final element of the concluding survey was to ask the users if they had any additional comments regarding the entire research study experience as a whole. Many users (25 of 52 - 48.1%) said that overall the study design was excellent, the experiment flow was comfortable, and they enjoyed participating. Six users wished the experimenter luck with the outcomes of the results. Eight users said they had fun and found the entire experiment interesting. Two users requested fewer saliva samples because they had difficulty producing enough spit. Four users wanted to hug a person and were disappointed they did not receive one. Five users mentioned that they did not appreciate the surprise public speaking and math tests or the associated stress that came with them. Three people wanted more information: one wanted why we gave the pre-visit instructions they had to follow, and two were interested in following up to hear the study results. One user requested fewer surveys and preferred them to be digital (they were all done on paper due to COVID). Finally, one user wished he could have one more hug before leaving.

7.6 DISCUSSION

In this section, we synthesize the results to evaluate the five hypotheses presented in Section 7.2. Afterward, we discuss the strengths and weaknesses of our robotic platform and possible modifications. Then, we discuss some limitations of our study design. Finally, we explore the potential implications this research has for the field of human-robot interaction.

7.6.1 *Hypothesis Testing*

This study used multiple measures, both self-reported and physiological, to get an accurate and holistic understanding of how the users felt at each time point. Thus, we go through each hypothesis and synthesize the data from each measure to draw a conclusion.

7.6.1.1 *[H1.] Any hug is better than no hug at all.*

QAS: While users in the no hug condition experienced an increase in negative emotions that approached significance from before to after the TSST, at no point did they experience a significant decrease in negative emotions, nor an increase in positive ones. Additionally, the only condition where users felt significantly less cared for was the no hug group. This group was also the only group to feel significantly less loved after the first intervention.

PANAS: While the no hug group did not experience any significant changes in positive emotions, all other hug conditions saw decreases in their self-reported positive emotions from their arrival to the first or second intervention or the debrief period. The no hug group had a significant increase in negative emotions after the TSST, as seen in both human hug groups. However, both the first and second intervention sessions and the debrief period significantly decreased the no hug group's negative emotions, without a significant difference observed between the two interventions. Both human hug groups saw the same significant decrease in negative emotions in both intervention sessions and the debrief period, while the robot hug group only saw a significant decrease after the first intervention, and the passive robot hug group saw no change. Thus, from the PANAS alone, we cannot determine that any hug is better than no hug at all.

Intervention Ratings From the intervention ratings, users in the no hug condition rated the robot hug as significantly more "enjoyable," "socially intelligent," and "friendly." Therefore, from the intervention ratings, we can see that *hugging a robot is significantly preferred over no hug.*

Re-Experience: In the first intervention round when users were split into five different groups, the "no hug" group had the lowest of all retention rates. Only 45.5% said that they would be willing to re-experience the empty room, whereas the other groups were 63.6% for the passive robot hug and robotic human hug group, 80% for the human hug, and 81.8% for the robot hug. Thus, from the re-experience ratings alone, *we can conclude that any hug is better than no hug at all for most people.*

Most Beneficial Session: When users were asked which session was the most beneficial for helping them alleviate their stress, only three out of eleven users said the no hug condition was the most beneficial. Therefore, from the most beneficial session question, *we can conclude that any hug is better than no hug at all.*

Spitting Time: The no hug group had significant decreases in the log of spitting time from arrival to post-intervention 1 (no hug), arrival to post recovery 1, and arrival to post-intervention 2 (robot hug). The passive robot hug, robot hug, and human hug groups did not see any significant differences. However, the robotic human group saw a significant decrease between arrival and post-intervention 1 and a decrease that approached significance between arrival and post-intervention 2. Thus, from spitting time data alone, we cannot conclude that any hug is better than no hug at all. Since this data was so variable, a larger participant pool must be sampled to draw a better conclusion.

Cortisol: All conditions saw significant decreases in cortisol levels between the TSST and recovery 2, while only the human hug group saw a significant decrease between the TSST and post-intervention 2. All groups saw significant decreases in cortisol between post-intervention 1 and recovery 1, post-intervention 1 and post-intervention 2, and post-intervention 1 and recovery 2. All conditions saw a significant decrease in cortisol between recovery 1 and post-intervention 2, and recovery 1 and recovery 2. Only the no hug group, passive robot hug group, and robot hug group saw significant decreases between the post-intervention 2 and recovery 2 periods. Additionally, the no hug group was the only group that saw an increase in cortisol between post wait and post-intervention 1. It seems that since in all other cases there was some form of hug intervention, this stopped the cortisol from rising, whereas in the no hug condition, there was nothing to help calm the users down, so their cortisol continued to rise until it reached a peak, and then gradually returned to normal over time. Therefore, from the cortisol data alone, *we can conclude that any hug is better than no hug at all.*

Oxytocin: No significant changes in the oxytocin levels were observed in the no hug group, passive robot hug, robot hug, or robotic human group. Thus, we cannot confirm nor deny that based on the oxytocin data alone, any hug is better than no hug at all. More data from more participants must be collected in order to make a determination.

User Comments and Experimenter Observations: While several users in the no hug group made their preference for solitude and silence clear, the majority expressed disappointment, confusion, and some even anger that we put them through such a stressful experience and then offered no help. When they later had the opportunity to hug the robot and then reflect on the two sessions, most users mentioned that it was better when they had someone or something to share their feelings with rather than being left alone in a room. The quality of words used to describe the room is also markedly different from those used to describe the intervention by users with any hug intervention. The no hug group described the room as, "productive," "boring," and "calming." Whereas the users in the hug conditions used words like, "comforting," "warm," "comfortable," "safe," "loved," and " peaceful." Additionally, users in the no hug condition spent the least amount of time of any group exploring and interacting with the environment. Therefore, using only the user comments and experimenter observations, *we can conclude that any hug is better than no hug at all*.

Conclusion: Synthesizing all these forms of data, we believe there is evidence to support that any hug is better than no hug at all.

7.6.1.2 *[H2.] A human hug is irreplaceable.*

QAS Only users in the human hug condition felt significantly more seen after an intervention method. Only users in the human hug condition felt significantly more cared for after the first intervention method. This group was also the only group to feel significantly more loved after the first intervention.

PANAS: The human hug group did not see an increase in positive emotions following their first intervention, which would help bolster this claim. However, no group saw an increase in positive emotions in any intervention sessions after the TSST. This group did experience significantly decreased negative emotions after the first and second intervention sesssions and the debrief period, just like the robotic human hug group and the no hug group. Therefore, from the PANAS alone, we cannot determine that a human hug is significantly better than the other intervention methods.

Intervention Ratings: Users in the human hug condition rated the human hug as significantly "safer," "more natural," and "friendlier." Therefore, from the intervention ratings we can see that *hugging a human behaving normally is significantly preferred over hugging any other agent.*

Re-Experience: After the first intervention, 80% of users in the human hug condition were willing to re-experience that intervention, whereas 81.8% of users in the robot hug condition were willing to re-experience that intervention. Therefore, from the intervention ratings, we *cannot conclude that hugging a human is irreplaceable.*

Most Beneficial Session: Of the ten users who experienced a human hug, only one user felt the robot hug was the more beneficial session. Therefore, from the most beneficial session question, *we can conclude that a human hug is irreplaceable.*

Spitting Time: The human hug group did not see any significant changes in spitting time across any of the seven time points, nor did either robot group. However, both the no hug group and the robotic human hug group did see changes. From spitting time data alone, we cannot determine whether the human hug is irreplaceable. Since this data was so variable, a larger participant pool must be sampled to draw a better conclusion.

Cortisol: Both the human hug and robot hug see significant decreases in cortisol between arrival and post-intervention 2, while the two robot conditions and two human conditions all have significant decreases between arrival and post-recovery 2. The human group has a decrease in cortisol between the post TSST and post-intervention 2 that the other groups do not have. However, the no hug group, the passive robot hug group, and the robot hug group, all have significant decreases in cortisol between the post-intervention 1 and post-recovery 2 that the human group do not have. All other changes in cortisol seem very similar between all the groups. Thus, from cortisol data alone, we cannot confirm or deny that a human hug is significantly better than any other hug.

Oxytocin: The human hug condition was the only condition in which we saw a significant increase in oxytocin. The increase was only measurable at the end of the experiment, and because it can take different amounts of time to show up in different users, we cannot determine if this increase was due to the human hug or the robot hug. Additionally, no significant changes in the oxytocin levels were observed in the no hug group, passive robot hug, robot hug, or robotic human group. Thus we cannot confirm nor deny that based on the oxytocin data alone, a human hug is irreplaceable. More data from more participants must be collected in order to make a determination.

User Comments and Experimenter Observations: Users were overall quite positive with their comments about the human hug condition. Users were more likely to say that the human hug reminded them of specific family members than in other conditions, though some users did say the robot hug and the robotic human hug reminded them of loved ones too. Half the users in this condition specially mentioned that they love hugs regardless of from whom they receive them. Additionally, the human hug condition was the only condition where several users (three) mentioned that they felt uncomfortable hugging a strange person. This objection did not show up for either the robot hug condition or the robotic human hug condition. Additionally, in the debrief forms, several users expressed disappointment that they did not get to hug another person, as it was mentioned in the advertisement material and informed consent document material that there was a possibility. Overall, more users had more positive comments to say about the human hug than other conditions. Based on only the user comments and experimenter observations, overall, users preferred the human hug condition the most.

Conclusion: Taking into account both the emotional and physiological data collected, we believe there is some evidence to support that a human hug is irreplaceable or at least slightly preferred over other intervention methods.

7.6.1.3 *[H3.] An active robot hug is preferred over a passive robot hug.*

QAS Users in the robot hug condition saw an increase in negative emotions around the TSST and a significant decrease after the second intervention period and the debrief, whereas the passive robot hug group did not see any change in their negative emotions. Therefore, using the QAS data, we do not have enough information to draw a conclusion. However, we do find evidence that an active robot hug can decrease negative emotions.

PANAS: Users in the active robot hug condition saw a significant decrease in positive emotions during the first intervention session, while users in the passive robot hug condition did not see a significant decrease until the second intervention session. However, the passive robot hug group did not significantly change their self-reported negative emotions, while the active robot hug group saw a significant decrease after the first intervention session. Thus, we cannot determine that an active robot hug is significantly better preferred over a passive robot hug from the PANAS alone.

Intervention Ratings: From the intervention ratings, no significant differences were observed in the perceived "safety," "naturalness," "enjoyability," "social intelligence," or "friendliness" between the active robot hug and the passive robot hug. However, users in the robot hug condition rated the second intervention with the robot as significantly more natural than the first interaction, thus indicating that *prolonged interaction with an active robot hug may be preferable to a passive robot hug.*

Re-Experience: After the first intervention, only 63.6% of users in the passive robot hug condition were willing to re-experience that intervention whereas 81.8% of users in the robot hug condition were willing to re-experience that intervention. Therefore, from the intervention ratings we can see that *hugging an active robot is preferred over hugging a passive robot.*

Most Beneficial Session: Of the ten users who experienced a passive robot hug, four users felt the passive robot hug was the more beneficial session, while six preferred the active robot hug. Therefore, from the most beneficial session question, *we can conclude an active robot hug is slightly preferred over a passive robot hug.*

Spitting Time: Neither the active robot hug group nor the passive robot hug group saw any changes in the log of spitting time between any of the seven time points collected. Thus, from spitting time data alone, we cannot draw any conclusions as to whether or not an active robot hug is better than a passive robot hug. In order to determine this, a larger participant pool must be sampled.

Cortisol: The robot hug group has a significant decrease in cortisol between arrival and post intervention 2 that the passive robot does not. All other changes in cortisol over the different time points are comparable. Looking at only the cortisol data, it appears that prolonged interaction with an active robot hug is more beneficial than hugging a passive robot.

Oxytocin: No significant changes in the oxytocin levels were observed in the passive robot hug or the robot hug group. Thus, we cannot confirm or deny that hugging an active robot is better than hugging a passive robot based on the oxytocin data alone. More data from more participants must be collected in order to make a determination.

User Comments and Experimenter Observations: The users in the passive robot hug condition mentioned in their comments that the robot did not seem very natural and was not a great hugger. While it was comfortable and warm, they mentioned that they would prefer to hug someone who could support them. In contrast, users in the robot hug group mentioned that they liked the feeling of the robot's embrace, occasional pro-active squeezes and appreciated the robot's ability to respond to the user actions. One petite user mentioned to the experimenter that she preferred hugging the robot in the passive hug condition because it did not hug as tightly, which was occasionally too tight for her liking during the squeeze. However, overall from the user comments, we believe there is support for the statement that an active robot hug is preferred over a passive robot hug.

Conclusion: While many data sources still require more participants to identify if any effects are present, by combining all different kinds of data we collected, we believe there is evidence to support that an active robot hug is preferred over a passive robot hug.

7.6.1.4 *[H4.] An active robot hug is preferred over a robotic human hug.*

QAS: Users in the robotic human hug condition saw an increase in positive emotions following the first intervention, whereas those in the robot hug condition did not. Users in the robot hug condition felt an increase in how cared for they felt after both active robot hug sessions, compared to after the TSST, while users in the robotic human hug did not experience a change. Conversely, the robotic human hug group felt a significant decrease in how in sync they felt between the first (robotic human) and second (robot) interventions, whereas there was no change for the robot hug group.

PANAS: Both the active robot hug and robotic human hug experienced significant decreases in positive emotions during the debriefing period compared to their arrival. While the active robot hug group saw a significant decrease after the first intervention and the robotic human hug saw a significant decrease after the second intervention, both these interventions represent the time the users interacted with the active robot hug. The robot hug group saw a decrease in self-reported negative emotions after the first intervention following the TSST. The robotic human hug group saw decreased negative emotions after the first and second intervention and the debriefing period. Thus, from only the PANAS, we can conclude that the robotic human hug was more beneficial to users than the active robot hug.

Intervention Ratings: Just as with the active versus passive robot hug, from the intervention ratings, no significant differences were observed in the perceived "safety," "naturalness," "enjoyability," "social intelligence," or "friendliness" between the active robot hug and the robotic human hug.

However, users in the robot hug condition rated the second intervention with the robot as significantly more natural than the first interaction, thus indicating that *prolonged interaction with an active robot hug may be preferable to a robotic human hug.*

Re-Experience: After the first intervention, only 63.6% of users in the robotic human hug condition were willing to re-experience that intervention whereas 81.8% of users in the robot hug condition were willing to re-experience that intervention. Therefore, from the intervention ratings, we can see that *hugging an active robot is preferred over hugging a robotic human.*

Most Beneficial Session: Of the ten users who experienced a robotic human hug, seven users thought the robotic human hug session was more beneficial, while only three users felt the robot hug session was beneficial. Therefore, from the most beneficial session question, *we can conclude that a robotic human hug was more beneficial than a delayed active robot hug.* However, when comparing the highest percentage of each group for the most beneficial session, the robot hug group had 72.72%, while the robotic human hug group only had a maximum of 70%. Therefore, from the most beneficial session question, *we determine that a robot hug is preferred over a robotic human hug.*

Spitting Time: The robotic human hug group saw significant decreases in the log of spitting time between arrival and post-intervention 1, indicating that perhaps they were less stressed after receiving the robotic human hug. The decrease in the log of spitting time between arrival and post-intervention 2 approached significance, also indicating that users were less stressed after hugging the robot. Users in the robot hug group saw no significant changes at any time point during the study. With such a small participant pool per condition, it is difficult to determine whether or not a robot hug or a robotic human hug is preferred based on spitting time data.

Cortisol: The robot hug group has a significant decrease in cortisol between arrival and post-intervention 2 that the robotic human hug does not exhibit. Additionally, the robot hug group has a significant decrease in cortisol between post-intervention 2 and recovery 2. All other changes in cortisol over the different time points are comparable. Looking at only the cortisol data, it appears that prolonged interaction with an active robot hug is more beneficial than hugging a robotic human.

Oxytocin: No significant changes in the oxytocin levels were observed in the active robot hug or robotic human group. We cannot confirm nor deny that based on the oxytocin data alone, any hug is better than no hug at all. More data from more participants must be collected in order to make a determination.

User Comments and Experimenter Observations: Users in both the robot hug and the robotic human hug condition were had positive comments to share about hugging their respective agents. In the robotic human hug group, users commented on the calming nature of the hugging agent's breath and the heartbeat, which were not present in the robot hug. However, users also commented that hugging a stranger was not comfortable or natural, whereas no one commented that they had not previously known the robot. In both conditions, users compared hugging the agent to hugging a loved one. In the second intervention session, users in the robot hug had increased positive comments, citing the increased trust from prolonged interaction with the robot. The users in the robotic human hug who experienced both hugging conditions said that the robotic human hug felt more "real," and the meaning behind the hug made them feel "loved and cared for." Based on user comments and experimenter observations, there is insufficient evidence to determine whether a robot hug or a robotic human hug is preferred.

Conclusion: While more data is still required for the physiological measures, from the data we do have, we can conclude that hugging an active robot is preferred over hugging a robotic human.

7.6.1.5 *[H5.] All users will benefit from a delayed active robot hug.*

QAS: Users in the human hug condition and robot hug condition saw decreases in negative emotions that approached significance and were significant, respectively, that only occurred after the second intervention (the delayed active robot hug). While users in the no hug condition were the only ones to feel significantly less cared for following the TSST and first intervention, the second intervention made them feel significantly more cared for, thus showing that even a delayed hug can be beneficial.

PANAS: The passive robot hug group, robot hug group, and robotic human hug group saw significant decreases in their self-reported positive emotions during their intervention method with the active robot hug (intervention 2 for passive robot hug and robotic human hug, intervention 1 for robot hug). The no hug group, the robot hug group, the robotic human hug group, and the human hug group all saw significant decreases in negative emotions during their intervention session with the active robot hug compared to the TSST. The passive robot hug group did not experience an increase in negative emotions during the TSST like the other four groups, and thus a significant decrease was not observed since it never peaked. Thus, we can conclude that most users benefited from a delayed active robot hug from only the PANAS.

Intervention Ratings: As discussed earlier, the no hug group found the delayed robot intervention significantly more "enjoyable," "socially intelligent," and "friendly," while the robot hug group found the delayed intervention significantly more "natural." The human hug group found the delayed robot hug significantly less "safe," "natural," and "friendly," while the robotic human hug did not report any significant difference. Therefore, from the intervention ratings alone, we *cannot conclude that all users benefited from a delayed robot hug. However, we can determine that the delayed active robot hug was beneficial for many.*

Re-Experience: Users' willingness to re-experience the intervention session in the no hug condition increased from 45.4% to 72.7% from the first to second intervention. In both the passive robot hug, robotic human hug, and the human hug conditions, users were just as willing to re-experience their first intervention method as they were the second (the delayed robot hug). One fewer person in the robot hug condition was willing to re-experience the intervention the second time, resulting in a decrease from 81.8% to 72.7%. Therefore, from the intervention ratings, we can conclude that *most users benefited from a delayed active robot hug.*

Most Beneficial Session: of the 52 users who participated, 26 users (exactly half) found the delayed robot hug more beneficial than any other earlier intervention. Thus, from the most beneficial session question, *we can conclude that half the users benefited from a delayed robot hug.*

Spitting Time: The no hug group saw a significant decrease between arrival and the post intervention 2 spitting time, while the robotic human hug group saw a decrease that approached significance between the arrival and post intervention 2 spitting time, thus indicating that users in both conditions felt less stressed after hugging the robot. Since the users in the robot hug, passive robot hug, and human hug groups saw no change in the log of spitting time during any point of the study, we cannot conclusively confirm or deny that all users benefited from a delayed robot hug. However, with a larger sample size, it does seem that the results are trending towards that direction.

Cortisol: The no hug group, passive robot hug group, and robot hug group all saw significant decreases in cortisol between post-intervention 2 and post-recovery 2. Though the decreases were not significant, the cortisol concentrations for the robotic human hug and the human hug between post-intervention 2 and recovery 2 are trending towards decreasing. With increased participant numbers, these changes will be more clear. Thus, from the current cortisol data available, we believe that it largely supports the hypothesis that all users benefit from a delayed robot hug.

Oxytocin: Once again, the human hug condition was the only condition in which we saw a significant increase in oxytocin. The increase was only measurable at the end of the experiment, and because it can take different amounts of time to show up in different users, we cannot determine if this increase was due to the human hug or the robot hug. Additionally, no significant changes in the oxytocin levels were observed in the no hug group, passive robot hug, robot hug, or robotic human group. Thus, we cannot confirm or deny that all users benefit from a delayed robot hug based on the oxytocin data alone. More data from more participants must be collected in order to make a determination.

User Comments and Experimenter Observations: Users had overwhelmingly positive things to say about the delayed robot hug during the second intervention session. Remarkably, several users even said that "hugging a robot wasn't much different from hugging another person." Users quickly spotted use cases where a hugging robot would come in handy during times when physical distance is required. Overall, they mentioned that the hugs, squeezes, and responsiveness was comfortable and pleasant. While some users may have been initially hesitant, they relaxed and calmed down after one or two hugs and found the hugs to be quite lovely and enjoyable. Several participants even pointed out that they preferred hugging the robot to a person because of the lack of social pressure and judgment. From user comments and experimenter observations, we can conclude that all users enjoyed and benefited from delayed robot hugs.

Conclusion: Combining all our sources of data, both emotional and physiological, we can conclude that most users enjoyed and benefited from delayed robot hugs.

7.6.2 *Strengths, Weaknesses, and Possible Modifications of the Robot*

7.6.2.1 *Strengths*

Overall, the fourth version of the HuggieBot platform is a suitable robot for close social-physical human-robot interactions with many strengths. HuggieBot's softness and warmth create an enjoyable tactile interface for users. The robot's size allows it to hug short and tall, small and large users with ease. The adaptive hugging modeled as a form of grasping ensures everyone receives a comfortable, custom, and secure embrace. The HuggieChest, a novel inflatable torso, allows for good detection of intra-hug gestures on the torso. The dual release detection (pressure and torque release) also works well to accommodate different user release styles and ensure no user is held in an embrace for too long.

Another strength is that the robot behaved in a very autonomous way. Though it is not quite yet ready for testing in-the-wild, it is certainly close. Users only watched a short instructional video and watched a live demonstration an hour and a half or two hours (depending on the session 1 intervention) before they hugged the robot. There were no practice hugs or tutorial sessions. Additionally, the experimenter was only in the room to record data and to physically intervene if necessary. The robot detected user presence, initiated a hug, and terminated the hug entirely autonomously.

The intra-hug gesture detection, classification, and response capabilities of HuggieBot are another major strength of this platform. To the best of our knowledge, no other hugging robot can detect or perform any such gestures. Based on the comments and ratings from our users, these intra-hug gestures significantly improve the overall quality and enjoyment of the hug.

7.6.2.2 *Weaknesses and Possible Modifications*

The robot's arms, the 6-DOF Kinova JACO arms, were selected for their slender form factor, quiet operation, and inherent safety features. One weakness of these arms is their low maximum speed, which is designed as part of the safety features. When users walk towards the robot, they must wait several seconds longer than they would for a human to finish enveloping them in a hug. In such an intimate social-physical interaction as a hug, several seconds can make a big difference in the perceived naturalness, and thus, overall enjoyability.

Another weakness related to the arms is that while we added padding on top of both Kinova JACO arms, it did not seem to be enough for some users. Therefore, we would add more padding to the arms for a more enjoyable tactile experience in the future. We also noticed that this complaint was more common if the user was particularly slender (categorized by BMI as "underweight.") Thus, we recommend in addition to increasing the robot's arm padding that we add to the robot's visual perception a rough estimation of the user's weight or BMI. Just as you would not squeeze a tiny child or a frail older person very tightly because their muscles are not as strong, so too do we think that robots should take extra caution and care when hugging those and other slender individuals who do not have as much padding on their own bodies.

As we learned from our previous work, users dislike a robot hug that releases them too soon equally as much as they dislike a robot that holds onto them long after they want to be released (Block & Kuchenbecker, 2019). Thus, it was necessary to ensure reliable release detection, so we use the robot's torso pressure and the joint torques on the robot's arms. Unfortunately, we did not anticipate that many users would release both hands off the robot's back to perform intra-hug gestures while still wanting the hug to continue. We thus found it very common for users to squeeze the robot tightly when they saw the robot releasing them from a hug in the hopes of getting it to hug them back again. Therefore, an essential modification of the robotic platform for the future will be to continue observing the torso pressure after the end of the hug has been initiated. If the user reapplies pressure, cancel the release of the hug and begin hugging the user again.

In this fourth version of the robot, we handled an edge case where if the robot does not observe a large enough increase in pressure to detect that a user had made contact with its back, it hugs the user and releases on a timed schedule. This method ensures that no users get stuck in an embrace and need to be released by the experimenter, which happened in previous experiments occasionally. Overall, this was a good addition. However, the duration of the hug was too short for many users' preferences. Therefore, increasing the overall duration of the timed hug and allowing the robot is a future modification that would increase the overall impression of the robot.

While the HuggieChest is an excellent strength of the robot as it allows for interactivity during the hug, it also is a slight weakness. The current HuggieChest is made of heat-sealed vinyl sheets, with additional HH-66 vinyl cement for extra sealing protection. Unfortunately, some users squeezed the robot so tightly that they popped holes in the torso, thus causing leaks. Leaks in the torso result in a deflated robot that is less enjoyable to hug and can also cause problems with both detections of intra-hug gestures and release detections. When the overall pressure decreases, the robot thinks the user wants to be released when he/she does not. Therefore, a critical future modification to improve the HuggieChest will be to find a sturdier heat-sealing method that can withstand stronger forces applied by users.

7.6.3 *Study Limitations*

While we succeeded in having an approximately equal gender balance, a roughly equal number of users in each condition, and a counter-balanced hug presentation order, several changes to the research study design might have improved it further. We had to disclose that there would be hugging of a robot and possibly a person to the users in the recruitment materials for COVID safety. Therefore, we had a self-selection bias in the users who chose to participate. Only those people who enjoy and want hugs signed up. Thus we did not have any users who refused to hug either the robot or the human. Additionally, based on the B2 level English requirement (due to the large number of surveys users were required to fill out), we also ended up with a highly educated participant population.

Additionally, while we did our best to recruit many participants, with only one experimenter and the long duration of the experiment (3.5 hours), only three users could participate in a single day. While a minimum of ten participants per condition is respectable, given that this study has five different conditions and that gender can affect oxytocin and cortisol levels, a greater number of users per condition would be preferable to detect true significant changes better.

Another limitation of our study was the oxytocin analysis. The lab that analyzed our samples did not lyophilize them, which has been known to remove contaminants from the samples before using an ELISA kit. Additionally, our samples were analyzed using an Enzo ELISA kit, which some experts believe is less sensitive and less reliable than the Arbor ELISA kit (Wirobski et al., 2021). We are currently looking into alternative analysis options for the oxytocin, including radio-immunoassay analysis to replace the initial Enzo ELISA results.

7.6.4 *Implications for HRI Research*

The interest in and positive feedback about HuggieBot 4.0 shows that there is a need for personal social-physical human-robot interaction. This research, in particular, highlights the importance of the sense of touch. We learned from this study that it matters less whom the touch comes from as long as the touch is comfortable, comforting, and firm. In all of our studies with HuggieBot, but this last one especially, we see that user opinion and perception of the robot improves after several touch interactions. Touch is surprisingly under-utilized in the field of human-robot interaction! Physical interaction should be integrated into other social robots the HRI community is already using for applications such as education, motivation, and exercise. Researchers may find users more receptive to instructions and feedback from the robot if it places its hand on the user's back or hand while it speaks to the user.

While everyone has their own opinions of style, our users also preferred when our robot was clothed similarly to a person. People told us the sweatshirt made the robot "look like a sports guy," and called the robot's mittens, "Bernie Sanders mittens." While we selected these items to add softness to the robot, the outfit helped the users laugh, relax, and feel more comfortable around the robot. Users particularly liked our timely addition of an FFP2 mask across the robot's face screen. In contrast, when we left the bottom half of the robot exposed in the online study in Chapter 5, users said the robot look, "creepy" and "unfinished." Thus, we recommend dressing up social robots rather than exposing a hard, metallic exterior. We also recommend taking note of and reflecting on any societal changes in how you present your robot.

Finally, an important contribution this work presented was expanding beyond only relying on surveys to supplement other quantitative data sources, and we urge the HRI community to do the same. In particular, we noticed that men typically under-report negative feelings, like those of stress or sadness, and typically over-report their capabilities. For example, in general, in this study, we found that men reported in the QAS that they did not feel very stressed by the TSST, while their heart rate and cortisol data suggest they did. In addition, we had the users self-report how well they think they performed in each section of the TSST, while we had our panel of experts also rate them. We found that male users, in general, reported they performed much better than they did when we averaged the three expert scores to create a "true performance" score. This disconnect tells us that using solely self-reported data is unreliable. Adding a few quantitative physiological measures to triangulate self-reported measures and free-response in other HRI studies can be illuminating.

7.7 CONCLUSION

This chapter outlined an exciting and essential step for HuggieBot. We ran a 3.5-hour user study with 57 participants collecting both emotional and physiological data. This fourth version of the robot is by far the most capable, enjoyable, and comfortable version yet. It used visual and haptic perception to adjust to the approach and size of the user autonomously. It also detected all user intra-hug gestures and responded with squeezes. With the integration of the timed release edge case, no users were ever stuck in an embrace if they applied insufficient pressure. This study also marks the first time any hugging robot has been compared to hugging another person, which is quite an ambitious goal. While we did not expect the robot to be preferred over a real human hug, the robot performed quite well and surpassed even the author's expectations. While it may look and feel different, the benefits and enjoyments users receive from hugging our robot are not so far from those derived from hugging another person with whom the user does not have a strong, personal connection.

8

CONCLUSION & OUTLOOK

I will not play tug o' war. I'd rather play hug o' war. Where everyone hugs instead of tugs, Where everyone giggles and rolls on the rug, Where everyone kisses, and everyone grins, and everyone cuddles, and everyone wins.

— Shel Silverstein, American Writer

In this thesis, we proposed HuggieBot as a human-sized interactive hugging robot with visual and haptic perception to supplement human hugs when obtaining that comfort from others is either difficult or impossible. We also presented eleven hugging tenets to which we believe future hugging robot designers should adhere during the creation of new robotic platforms for close, personal, social-physical human-robot interactions. Over the course of this dissertation, both the robot and the study designs became more sophisticated.

In Chapter 4 we tested and validated the first two tenets, that a robot should be soft and warm, along with the first version of the robot, a Willow Garage Personal Robot 2 (PR2) with software and hardware upgrades. We found that users enjoy exchanging hugs with a robot but found the physical platform unsatisfactory for this intimate social-physical human-robot interaction. However, we did find enough support for the idea to merit further exploration.

In Chapter 5 we presented the second version of HuggieBot, which, other than two Kinova JACO arms, was a completely custom robotic platform. We designed and built this robot according to the first six hugging robot tenets we also presented. The first two tenets were those tested in Chapter 4, while the final four were new. We iterated the design of our robot over two user studies (one online and one in person). In the first study, we established user preference for our new robot over the old version, and in the second study, we found evidence to support five of our proposed six tenets.

Chapter 6 featured two user studies. In the first study, we collected data from 32 participants to create a detection and classification algorithm for intra-hug gestures. We also gathered user preferences for our probabilistic behavior algorithm on how the robot should respond to these gestures. We validated these algorithms in the second study with 16 users and the final six hug tenets, for a total of eleven.

In Chapter 7 we tested the most sophisticated version of our robotic platform. The robot behaved in a completely autonomous way and was directly compared to hugging a friendly but unfamiliar person. We collected both emotional (self-reported) and physiological (quantitative) data for the entire 3.5 hours that each of our 57 participants was in the study. We found users pleasantly surprised by the quality of active hugs that HuggieBot could deliver, with many users commenting that there was actually not much difference between the robot and the human hug.

8.1 PUBLIC INTEREST

A unique and exciting element of this thesis is that this project seems to have captivated the interest and imagination of the public. We have been fortunate that this project has received significant and consistent press coverage throughout all versions of the robot.

[Upcoming]	National Geographic has conducted an interview with Alexis and Katherine, and is planning a trip to film the journalist hugging HuggieBot.
21 July 2021	Women's Wealth recorded a podcast interview with Alexis about HuggieBot, and it is now available on Spotify and Apple Podcasts!
8 July 2021	The Current (Canadian Broadcasting Corporation) did a story about hugging and interviewed Alexis to learn about HuggieBot, why she made it, and how it would succeed after the pandemic.
24 June 2021	The Globe (ETH Magazine) featured Alexis and HuggieBot in the three page article "Robots for comfort and counsel" in the second issue of the magazine in 2021 focusing on "Nature and design: blurring the boundaries between the natural and the artificial"
10 March 2021	Entrepreneur.com published an article titled "What to

	Do If People Laugh at Your Idea" about Alexis E. Block and HuggieBot 3.0
11 February 2021	TechXplore published an article titled "HuggieBot 2.0: A Soft and Human-Size Robot That Hugs Users on Request" with an interview with Alexis focusing on the Six Hug Commandments HRI 2021 Paper
16 January 2021	SWR TV (German TV News Station) filmed and aired an interview with Alexis E. Block, Katherine J. Kuchenbecker and a live demo of HuggieBot 3.0
December 2020	PM Magazin (German Science Magazine) published an article titled "Touched by a Robot" about Alexis E. Block and HuggieBot 2.0
16 October 2020	The New York Times published an article titled "When We Can Hug Again, Will We Remember How It Works?" featuring with an interview with Alexis and discussion of HuggieBot
15 October 2020	Double Helix (Australian Science Magazine) published an article titled "The Comfort of Robot Hugs" about Alexis E. Block and HuggieBot 2.0
6 December 2019:	Stuttgarter Zeitung published an article with an interview of Alexis E. Block's research and focused on the new version of HuggieBot 2.0, in German: "Ein Roboter spendet Trost"
27 May 2019:	The Robot Report published an article with an interview about Alexis E. Block's research and recent invited talk: "ETH Zurich Researcher Works to Build Human-Machine Trust, One Robotic Hug at a Time"
26 May 2019:	STEM On Fire interviewed Alexis E. Block for their podcast to encourage high-school juniors and seniors and college freshmen and sophomores to study STEM
30 August 2018:	NowThis Future Media created a short video using experimental footage and a summary of the experimental

procedure and results published in Alexis' HRI Pioneers paper, the video reached over 250,000 views

16 June 2018: NPR (National Public Radio) HuggieBot was featured as a question on NPR's "Wait Wait Don't Tell Me" game show

15 June 2018: The Paul Ross Show did an 11 minute interview with Alexis E. Block that was broadcast on TalkRadio in the UK

12 June 2018: The Times (UK Newspaper) published an article written about HuggieBot featuring an interview with Alexis E. Block: "Feel the Love with a Robo-Hug That's Better Than the Real Thing"

11 June 2018: NBC News published an article written about HuggieBot featuring an interview with Alexis E. Block: "Why Scientists are Teaching this Burly Robot to Hug"

7 June 2018: Digital Trends published an article written about HuggieBot featuring an interview with Alexis E. Block: "Forget Roomba, Your Most Important House Robot Could be the One that Hugs You"

5 June 2018: IEEE Spectrum published a long article written about HuggieBot featuring an interview with Alexis E. Block: "The Importance of Teaching Robots to Hug"

8.2 FUTURE WORK

While there are several possibilities for future work related to this thesis, we are most interested in two specific directions. First, several data streams from the final study remain to be analyzed. Second, we hope to create a future system to deliver good hugs in a more natural setting, without the need for any instruction, demonstration, or practice trials.

8.2.1 *Additional Data Streams and Analysis from Emotional and Physiological Effect of Robot Hugs Study*

The final study of this thesis captured so many different data streams that we could not yet investigate all thoroughly.

We have collected continuous heart rate data for all participants throughout the 3.5-hour user study. We are interested in using this data to calculate heart rate variability (HRV), which can be a good indication of actual stress and relaxation (Kaur et al., 2014; Kim et al., 2018; Sarang & Telles, 2006). While heart rate focuses on the average beats per minute, heart rate variability measures the specific changes in time (or variability) between successive heartbeats. An increase in heart rate and a decrease in heart rate variability indicate an adverse effect of the autonomic nervous system. We are also interested in comparing each user's self-reported emotional state (like stress) with their heart rate and heart rate variability to see how well the self-report and physiological data correspond and whether there are any significant differences between genders; the investigator observed anecdotal evidence in support of this hypothesis during the study.

While we collected user self-report and the panel of expert ratings on each participant's performance in the TSST, analyzing this data was out of the scope of this thesis. We are interested in looking into this further to see if the trend observed in Chapter 7 Section 7.5.5.3 of men over-estimating and women under-estimating their performance holds true across participants.

Another data stream we collected that we are currently investigating is the video footage of the two interventions. Each of the hug rooms (human and robot) was instrumented with three cameras: one behind the hugging agent, one behind the user, and one side view pointed directly at the hug. We are looking to see if the total duration of the hug affects user heart rate, oxytocin, and cortisol. As we kept users in the room for ten minutes but allowed them to hug for as long or as short a time as they wanted, some users had more time exposed to the beneficial deep pressure touch therapy than others, which could account for the small number of significant differences in oxytocin we observed.

Another aspect of the hugging videos we are exploring is how close users got to the hugging agent during their embrace. Some users kept distance between their chests; some users allowed their chests to touch but kept their lower halves angled away, while still others got very close and snuggled up to the hugging agent. We are quantifying these elements through video coding and plan to investigate whether these aspects affect the heart rate, oxytocin, and cortisol of users.

We are interested about conducting a second, more precise analysis of our saliva samples for oxytocin. As mentioned in Chapter 7 Section 7.6.3, we have several concerns regarding the oxytocin analysis. The samples were not lyophilized, and were not analyzed using the most sensitive or reliable methods. Research has shown that radio-immunoassay analysis is a promising and potentially more reliable method of analysis than using an enzyme immunoassay, like the ELISA kits (Szeto et al., 2011). We have identified a potential lab to conduct radio-immunoassay analysis of our samples. We are curious to compare these results with the initial Enzo ELISA results.

As mentioned earlier, we are also interested in investigating whether there is any correlation between spitting time, cortisol, and or oxytocin. We are particularly interested in investigating whether the preliminary trends we identified in Chapter 7 Section 7.5.4.5 regarding a potential correlation of spitting time and cortisol holds true. We are also interested in looking at the interaction between cortisol and oxytocin, which is still somewhat of a mystery to psychologists. In general, cortisol and oxytocin interact and help regulate one other; however, this interaction presents itself quite differently between men and women. In men, the hormone vasopressin typically takes over and overpowers oxytocin during a cortisol spike. However, for women, oxytocin surges and floods the body to accommodate the additional stress or pain. More participants would need to be recruited to investigate this relationship further, and a thorough analysis must be conducted.

In future work, we are looking forward to exploring these additional data streams more systematically.

8.2.2 *New Research Avenues to Explore*

Once the COVID-19 crisis has ended, we believe it would be beneficial to conduct a thorough in-the-wild study of this kind of hugging robot. Specifically, we envision deploying a future version of HuggieBot in a pedestrian environment to see how many everyday people would and would not be interested in hugging a robot. To do this, we will need to improve the robot's approaching user detection algorithm, which currently works well only with one user in the frame. To function in a non-controlled (non-lab) environment with many passersby potentially in the robot's field of view, we will need to make this aspect of the robot more robust.

HuggieBot 4.0 can detect if a user changes their mind and decides not to hug the robot. Currently, it stays with its arms up in case the user decides to come back. We will need to detect if a user walks away and then have the robot lower its arms and wait for a new user to approach. Another challenge of conducting an in-the-wild study will be dealing with non-haptic noise detected by the microphone. We are currently taking a baseline at the start of every hug to determine the mean microphone output, but we are not estimating the magnitude or spectrum of background noise that needs to be filtered out. We found that nearby construction during Chapter 6 sometimes caused high levels of ambient noise during pilot testing, causing our perception pipeline to mistakenly think that the user was continuously rubbing the robot's back.

A simpler improvement to the robot's behavior centers on reducing the likelihood that users accidentally end the hug before they intend to do so. We noticed that several users in our studies briefly took their hands off the robot's torso to adjust their grip and grasp the robot tighter when performing a squeeze. Because the chamber pressure decreased, the robot assumed the user wanted to be released, ending the hug. To make our system more robust, rather than act on a single low-pressure value, if the pressure decreases and increases again quickly, the robot should not release the user because they probably do not want the hug to end. Given the negative social impact of not letting go of a user who wants to be released, finding the right balance between fast and reliable hug termination and avoidance of accidental user releases is important.

A third element we are looking forward to researching is the extent to which a future version of HuggieBot can help strengthen personal relationships between people who are separated by a physical distance. To do this, we have already developed a mobile app, the HuggieApp, that allows remote users to send customized hugs to local users (via the robot). The sender can customize the hug's duration and tightness, and they can add a variety of intra-hug gestures. They can even replace the robot's animated face with a video message for the receiver. The local user redeems the hug by scanning a custom QR code on their mobile phone at the robot's camera, which is located above its face screen. Users can re-redeem their favorite hugs as many times as they want, which could be especially meaningful if the original sender has passed away. For several months, we would propose to observe pairs of platonic users interacting with each other through this future version of HuggieBot. We believe it would be interesting to evaluate the relationship's perceived closeness to see the extent to which this type of embodied affective robot can help bridge physical distance between people.

A final area we are interested in investigating is enabling this robot to work as a diagnostic or evaluation tool. We are curious about using machine learning to equip the robot with the ability to sense a person's mood and perhaps their mental health status based on conversations, facial expressions, an emotion tracking device, and the way a person hugs the robot, which is a form of haptic intelligence that only people possess at the present.

The applications for this research are widespread. Immediate uses could be on college campuses to help alleviate student stress or nursing homes to try to bring happiness to the residents. As the research progresses, the platform could be adapted for use in rehabilitation centers or for children with autism. A final area for implementation for this research could be through co-parenting using telepresence robots (Neustaedter & Yang, 2017), to enable parents and children to connect in real time when far away.

BIBLIOGRAPHY

About Adult BMI [Online; accessed 10-June-2021]. (2020). https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html

Ainsworth, M. D. S. (1982). 14. the development of infant-mother attachment. *In thebeginning* (pp. 133–143). Columbia University Press.

Angelini, L., Caon, M., Lalanne, D., Khaled, O. A., & Mugellini, E. (2015). Towards an anthropomorphic lamp for affective interaction. *Proceedings of the International Conference on Tangible, Embedded, and Embodied Interaction (TEI)*, 661.

Anzalone, S. M., Varni, G., Ivaldi, S., & Chetouani, M. (2017). Automated prediction of extraversion during human-humanoid interaction. *International Journal of Social Robotics*, 385.

Arns, M., Laliberte, T., & Gosselin, C. (2017). Design, control and experimental validation of a haptic robotic hand performing human-robot handshake with human-like agility. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, 4626.

Baisch, S., Kolling, T., Schall, A., Rühl, S., Selic, S., Kim, Z., Rossberg, H., Klein, B., Pantel, J., Oswald, F., & Knopf, M. (2017). Acceptance of social robots by elder people: Does psychosocial functioning matter? *International Journal of Social Robotics*, 293.

Bales, E., Li, K. A., & Griwsold, W. (2011). CoupleVIBE: Mobile implicit communication to improve awareness for (long-distance) couples. *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)*, 65.

Barber, J., Volz, R. A., Desai, R., Rubinfeld, R., Schipper, B., & Wolter, J. (1986). Automatic two-fingered grip selection. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, *3*, 890.

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. *Fundamentals of Nonverbal Behavior*, 401.

Birkett, M. A. (2011). The Trier Social Stress Test protocol for inducing psychological stress. *Journal of Visualized Experiments*, (56), 1.

Block, A. E. (2017). *How should robots hug?* (Master's thesis). University of Pennsylvania - School of Engineering and Applied Science. Robotics.

Block, A. E., Christen, S., Gassert, R., Hilliges, O., & Kuchenbecker, K. J. (2021). The six hug commandments: Design and evaluation of a human-sized hugging robot with visual and haptic perception. *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 380.

Block, A. E., & Kuchenbecker, K. J. (2018). Emotionally supporting humans through robot hugs. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 293.

Block, A. E., & Kuchenbecker, K. J. (2019). Softness, warmth, and responsiveness improve robot hugs. *International Journal of Social Robotics*, *11*(1), 49.

Block, A. E., Seifi, H., Hilliges, O., Gassert, R., & Kuchenbecker, K. J. (2021). In the arms of a robot: Designing autonomous hugging robots with intra-hug gestures [under review]. *ACM Transactions on Human-Robot Interaction (THRI) Special Issue on Designing the Robot Body: Critical Perspectives on Affective Embodied Interaction*, 1.

Block, A. E., Young Kuchenbecker, S., Gassert, R., Hilliges, O., & Kuchenbecker, K. J. (2021). Love, actually? evaluating the emotional and physiological effects of robot hugs [in preparation]. *Journal to be determined*.

Brščić, D., Kidokoro, H., Suehiro, Y., & Kanda, T. (2015). Escaping from children's abuse of social robots. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 59.

Brummett, B., Boyle, S., Kuhn, C., Siegler, I., & Williams, R. (2009). Positive affect is associated with cardiovascular reactivity, norepinephrine level, and morning rise in salivary cortisol. *Psychophysiology*, *46*(4), 862.

Burns, R. B., Seifi, H., Lee, H., & Kuchenbecker, K. J. (2021). Getting in touch with children with autism: Specialist guidelines for a touch-perceiving robot. *Paladyn, Journal of Behavioral Robotics*, *12*, 115.

Campbell, J., & Yamane, K. (2020). Learning whole-body human-robot haptic interaction in social contexts. *CoRR*, *abs/2005.12508*.

Carter, C. S., Kenkel, W. M., MacLean, E. L., Wilson, S. R., Perkeybile, A. M., Yee, J. R., Ferris, C. F., Nazarloo, H. P., Porges, S. W., Davis, J. M., et al. (2020). Is oxytocin "nature's medicine"? *Pharmacological Reviews*, *72*(4), 829.

Cascio, C. J., Moore, D., & McGlone, F. (2019). Social touch and human development. *Developmental Cognitive Neuroscience*, *35*, 5.

Chang, J., MacLean, K., & Yohanan, S. (2010). Gesture recognition in the Haptic Creature. In A. M. L. Kappers, J. B. F. van Erp, W. M. Bergmann Tiest, & F. C. T. van der Helm (Eds.), *Haptics: Generating and perceiving tangible sensations. eurohaptics 2010* (pp. 385–391). Springer Berlin Heidelberg.

Chang, W. L., Šabanović, S., & Huber, L. (2013). Use of seal-like robot PARO in sensory group therapy for older adults with dementia. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 101.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods between subject and within subject design. *Journal of Economic Behavior and Organization*, *81*, 1.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893.

Chen, T. L., King, C.-H. A., Thomaz, A. L., & Kemp, C. C. (2014). An investigation of responses to robot-initiated touch in a nursing context. *International Journal of Social Robotics*, 141.

Chen, Y., Yu, M., Bruck, H. A., & Smela, E. (2016). Stretchable touch-sensing skin over padding for co-robots. *Smart Materials and Structures*, *25*(5), 1.

Choi, W., Pantofaru, C., & Savarese, S. (2011). Detecting and tracking people using an RGB-D camera via multiple detector fusion. *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 1076.

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.

Cooney, M. D., Becker-Asano, C., Kanda, T., Alissandrakis, A., & Ishiguro, H. (2010). Full-body gesture recognition using interial sensors for playful interaction with small humanoid robot. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2276.

Cooney, M. D., Nishio, S., & Ishiguro, H. (2012). Recognizing affection for a touch-based interaction with a humanoid robot. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1420.

Costanzo, M., De Maria, G., & Natale, C. (2020). Two-fingered in-hand object handling based on force/tactile feedback. *IEEE Transactions on Robotics*, *36*(1), 157.

Cramer, H., Kemper, N., Amin, A., Wielinga, B., & Evers, V. (2009). 'Give me a Hug': The effects of touch and autonomy on people's responses to embodied social agents. *Computer Animation and Virtual Worlds*, *20*(437-445).

Dalal, N., Triggs, B., & Schmid, C. (2006). Human detection using oriented histograms of flow and appearance. *Proceedings of the European Conference on Computer Vision (ECCV) - Volume Part II*, 428.

de Winter, J. C. F., & Dodou, D. (2017). *Experimental design in: Human subject research for engineers*. SpringerBriefs in Applied Sciences; Technology.

de Waal, F. (2019). *Mama's last hug: Animal emotions and what they tell us about ourselves*. W. W. Norton & Company.

DiSalvo, C., Gemperle, F., Forlizzi, J., & Montgomery, E. (2003). The Hug: An exploration of robotic form for intimate communication. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 403.

Ditzen, B., Neumann, I. D., Bodenmann, G., von Dawans, B., Turner, R. A., Ehlert, U., & Heinrichs, M. (2007a). Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology*, *32*(5), 565.

Ditzen, B., Neumann, I. D., Bodenmann, G., von Dawans, B., Turner, R. A., Ehlert, U., & Heinrichs, M. (2007b). Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology*, *32*, 565.

Duvall, J. C., Dunne, L. E., Schleif, N., & Holschuh, B. (2016). Active hugging vest for deep touch pressure therapy. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, 458.

Eckstein, M., Mamaev, I., Ditzen, B., & Sailer, U. (2020). Calming effects of touch in human, animal, and robotic interaction - scientific state-of-the-art and technical advances. *Frontiers in Psychiatry*, *11*, 1153.

Edelson, S. M., Goldberg, M., Kerr, D. C. R., & Grandin, T. (1999). Behavioral and physiological effects of deep pressure on children with autism: A pilot study evaluating the efficacy of grandin's hug machine. *53*(1979), 145.

Feil-Seifer, D., & Matarić, M. (2005). Defining socially assistive robotics. *Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics, 2005*, 465.

Feldman, R. (2017). The neurobiology of human attachments. *Trends in cognitive sciences*, *21*(2), 80.

Feldman, R. (2020). What is resilience: An affiliative neuroscience approach. *World Psychiatry*, *19*(2), 132.

Feldman, R., Rosenthal, Z., & Eidelman, A. I. (2014). Maternal-preterm skin-to-skin contact enhances child physiologic organization and cognitive control across the first 10 years of life. *Biological psychiatry*, *75*(1), 56.

Filipowicz, L., & Keller, J. (2018). How to send someone your heartbeat with Apple Watch or iPhone.

Fitter, N. T., Hawkes, D. T., & Kuchenbecker, K. J. (2016). Rhythmic timing in playful human-robot social motor coordination. *Proceedings of the International Conference on Social Robotics*, 296.

Fitter, N. T., & Kuchenbecker, K. J. (2016a). Designing and assessing expressive open-source faces for the baxter robot. *Proceedings of the International Conference on Social Robotics*, 340.

Fitter, N. T., & Kuchenbecker, K. J. (2016b). Equipping the Baxter robot with human-inspired hand-clapping skills. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication, (RO-MAN)*, (Section IV), 105.

Fitter, N. T., & Kuchenbecker, K. J. (2019). How does it feel to clap hands with a robot? *International Journal of Social Robotics*, *12*(1), 113.

Fitter, N. T., Mohan, M., Kuchenbecker, K. J., & Johnson, M. J. (2020). Exercising with Baxter: Preliminary support for assistive social-physical human-robot interaction. *Journal of NeuroEngineering and Rehabilitation*, *17*(19).

Fitter, N. T., Strait, M., Bisbee, E., Mataric, M. J., & Takayama, L. (2021). You're wigging me out! is personalization of telepresence robots strictly positive?

Floyd, K. (2000). Affectionate same-sex touch: The influence of homophobia on observers' perceptions. *The Journal of social psychology*, *140*(6), 774.

Fosshage, J. L. (2000). The meaning of touch in psychoanalysis: A time for reassessment. *Psychoanalytic Inquiry*, *20*(1), 21.

Friendship Lamps [Online; accessed 21-March-2021]. (2020). https://www.friendlamps.com

Garrell, A., Villamizar, M., Moreno-Noguer, F., & Sanfeliu, A. (2017). Teaching robot's proactive behavior using human assistance. *International Journal of Social Robotics*, 231.

Genz, T. (2007). Wearable haptic telecommunication device and system.

Glas, D. F., Wada, K., Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2017). Personal greetings: Personalizing robot utterances based on novelty of observed behavior. *International Journal of Social Robotics*, 181.

Grandin, T. (1992). Calming effects of deep touch pressure in patients with autistic disorder, college students, and animals. *Journals of Child and Adolescent Psychopharmacology*, *2*(1), 63.

Grant, D. A. (1948). The Latin square principle in the design and analysis of psychological experiments. *Psychological Bulletin*, *45*(5), 427.

Grewen, K., Girdler, S., Amico, J., & Light, K. (2005). Effects of partner support on resting oxytocin, cortisol, norepinephrine, and blood pressure before and after warm partner contact. *Psychosomatic Medicine*, *67*, 531.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Jr., A. R. D., Durbin, M., Edmonson, M. S., Fischer, J. L., Hymes, D., Kimball, S. T., Barre, W. L., Lynch, F., J, S., McClellan, J. E., Marshall, D. S., Milner, G. B., Sarles, H. B., Trager, G. L., & Vayda, A. P. (1968). Proxemics [and Comments and Replies]. *Current Anthropology*, *9*(2/3), 83.

Harlow, H. F., & Zimmerman, R. R. (1959). Affectional responses in the infant monkey. *Science*, *130*(373), 421.

Hedayati, H., Bhaduri, S., Sumner, T., Szafir, D., & Gross, M. D. (2019). HugBot: A soft robot designed to give human-like hugs. *Proceedings of the ACM International Conference on Interaction Design and Children*, 556.

Heerink, M., Krose, B., Evers, V., & Wielinga, B. J. (2009). Measuring acceptance of an assistive social robot: A suggested toolkit. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 528.

Hey Bracelet [Online; accessed 4-April-2021]. (2020). https://www.feelhey.com

Holm, S. (1978). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65.

Hudson, J., Orviska, M., & Hunady, J. (2017). People's attitudes to robots in caring for the elderly. *International Journal of Social Robotics*, 199.

Insel, T. R. (1992). Oxytocin—a neuropeptide for affiliation: Evidence from behavioral, receptor autoradiographic, and comparative studies. *Psychoneuroendocrinology*, *17*(1), 3.

Juruena, M. F., Eror, F., Cleare, A. J., & Young, A. H. (2020). The role of early life stress in hpa axis and anxiety. *Anxiety Disorders*, 141.

Kaplish, A., & Yamane, K. (2019). Motion retargeting and control for teleoperated physical human-robot interaction. *19th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2019, Toronto, ON, Canada, October 15-17, 2019*, 723.

Kaur, B., Durek, J. J., O'Kane, B. L., Tran, N., Moses, S., Luthra, M., & Ikonomidou, V. N. (2014). Heart rate variability (hrv): An indicator of stress. *Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII*, *9118*, 91180V.

Keverne, E. B., & Kendrick, K. M. (1992). Oxytocin facilitation of maternal behavior in sheep a. *Annals of the New York Academy of Sciences*, *652*(1), 83.

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.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'trier social stress test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76.

Kory, J., & Breazeal, C. (2014). Storytelling with robots: Learning companions for preschool children's language development. *The 23rd IEEE international symposium on robot and human interactive communication*, 643.

Kosuge, K., Takeda, T., Hirata, Y., Endo, M., Nomura, M., Sakai, K., Koizumi, M., & Oconogi, T. (2008). Patner Ballroom Dance Robot - PBDR. *SICE Journal of Control, Measurement, and System Integration*, *1*(1), 74.

Krauss, K. E. (1987). The effects of deep pressure touch on anxiety. *The American Journal of Occupational Therapy*, *41*(6), 366.

Kubzansky, L. D., Mendes, W. B., Appleton, A. A., Block, J., & Adler, G. K. (2012). A heartfelt response: Oxytocin effects on response to social stress in men and women. *Biological Psychology*, *90*(1), 1.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. HarperCollins.

Laplante, J., Michaud, F., Larouche, H., Duquette, A., Caron, S., Letourneau, D., & Masson, P. (2005). Autonomous spherical mobile robotic to study child development. *IEEE Trans. on Systems, Man, and Cybernetics*, *10*.

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 [Current Trends in Women's Health Research]. *Biological Psychology*, *69*(1), 5.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *Lecture Notes in Computer Science*, 21.

Lynott, D., Corker, K. S., Wortman, J., Connell, L., Donnellan, M. B., Lucas, R. E., & O'Brien, K. (2014). Replication of "experiencing physical warmth promotes interpersonal warmth" by Williams and Bargh (2008). *Social Psychology*, *45*(3), 216.

Ma, R. R., Spiers, A., & Dollar, A. M. (2016). M2 gripper: Extending the dexterity of a simple, underactuated gripper. *Advances in Reconfigurable Mechanisms and Robots II. Mechanisms and Machine Science*, *36*.

Mahalski, P. A. (1981). The incidence of attachment objects and oral habits at bedtime in two longitudinal samples of children aged 1.5-7 years. *Journal of Child Psychology and Psychiatry*, *24*(2), 283.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71.

May, D. C., Holler, K. J., Bethel, C. L., Strawderman, L., Carruth, D. W., & Usher, J. M. (2017). Survey of factors for the prediction of human comfort with a non-anthropomorphic robot in public spaces. *International Journal of Social Robotics*, 165.

Mead, R. (2016). *Situated proxemics and multimodal communication: Space, speech, and gesture in human-robot interaction* (Doctoral dissertation). University of Southern California.

Mead, R., & Matarić, M. J. (2015). Robots have needs too: People adapt their proxemic preferences to improve autonomous robot recognition of human social signals. *New Frontiers in Human-Robot Interaction*, *100*, 100.

Mead, R., & Matarić, M. J. (2017). Autonomous human–robot proxemics: Socially aware navigation based on interaction potential. *Autonomous Robots*, *41*(5), 1189.

Meaney, M. J. (2001). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual review of neuroscience*, *24*(1), 1161.

Michaud, F., & Caron, S. (2002). Roball, the rolling robot. *Autonomous robots*, *12*(2), 211.

Michaud, F., Salter, T., Duquette, A., Mercier, H., Lauria, M., Larouche, H., & Larose, F. (2007). Assistive technologies and child-robot interaction. *AAAI spring symposium on multidisciplinary collaboration for socially assistive robotics*.

Miyashita, T., & Ishiguro, H. (2004). Human-like natural behavior generation based on involuntary motions for humanoid robots. *Robotics and Autonomous Systems*, *48*(4), 203.

Morrison, C. M., & Gore, H. (2010). The relationship between excessive Internet use and depression: A questionnaire-based study of 1,319 young people and adults. *Psychopathology*, *43*(2), 121.

Mummolo, J., & Peterson, E. (2017). Demand effects in survey experiments: An empirical assessment. *doi:10.2139/ ssrn.2956147*.

Nagy, E. (2011). Sharing the moment: The duration of embraces in humans. *Journal of Ethology*, *29*(2), 389.

Nakanishi, J., Sumioka, H., & Ishiguro, H. (2020). Mediated hugs modulate impressions of hearsay information. *Advanced Robotics*, *34*(12), 781.

Neira, C. J. B., & Barber, B. L. (2014). Social networking site use: Linked to adolescents' social self-concept, self-esteem, and depressed mood. *Australian Journal of Psychology*, *66*(1), 56.

Neustaedter, C., & Yang, L. (2017). Family communication over distance through telepresence robots. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW) Workshop on Robots in Groups and Teams*.

Nie, J., Park, M., Marin, A. L., & Sundar, S. S. (2012). Can you hold my hand?: Physical warmth in human-robot iinteraction. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, 201.

Nomura, T., Kanda, T., Kidokoro, H., Suehiro, Y., & Yamada, S. (2016). Why do children abuse robots? *Interaction Studies*, *17*(3), 347.

Nunez, E., Hirokawa, M., Igarashi, H., & Suzuki, K. (2017). Robotic device to mediate human-human hug-driven remote communication. *Lecture Notes in Artificial Intelligence (LNAI)*, *10652*, 198.

Odabasi, C. (2017). ROS people object detection and action recognition Tensorflow.

Oren, M., Papageorgiou, C., Sinha, P., Osuna, E., & Poggio, T. (1997). Pedestrian detection using wavelet templates. *Proceedings of IEEE Com-*

puter Society Conference on Computer Vision and Pattern Recognition, 193.

Pakanen, M., Colley, A., Häkkilä, J., Kildal, J., & Lantz, V. (2014). Squeezy bracelet – designing a wearable communication device for tactile interaction. *Proceedings of the NordiCHI: The Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, 305.

Park, E., & Lee, J. (2014). I am a warm robot: The effects of temperature in physical human-robot interaction. *Robotica*, *32*, 133.

Park, H. W., Gelsomini, M., Lee, J. J., & Breazeal, C. (2017). Telling stories to robots: The effect of backchanneling on a child's storytelling. *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, 100.

Paul, M., Haque, S., & Chakraborty, S. (2013). Human detection in surveillance videos and its applications: A review. *Eurasip Journal on Applied Signal Processing*, *176*, 1.

Peng, H., Zhou, C., Hu, H., Chao, F., & Li, J. (2015). Robotic dance in social robotics - a taxonomy. *IEEE Transactions on Human-Machine Systems*, *45*(3), 281.

Perry, B. D., & Szalavitz, M. (2006). *The boy who was raised as a dog: And other stories from a child psychiatrist's notebook-what traumatized children can teach us about loss, love, and healing*. Basic Books New York.

Pino, M., Boulay, M., Jouen, F., & Rigaud, A. (2015). "are we ready for robots that care for us?" attitudes and opinions of older adults toward socially assistive robots. *Frontiers in Aging Neuroscience*, *7*, 141.

Pulido, J. C., González, J. C., Suárez-Mejías, C., Bandera, A., Bustos, P., & Fernández, F. (2017). Evaluating the child-robot interaction of the NAOTherapist platform in pediatric rehabilitation. *International Journal of Social Robotics*, 343.

Rai, B., Kaur, J., Anand, S., & Jacobs, R. (2011). Salivary stress markers, stress, and periodontitis: A pilot study. *Journal of periodontology*, *82*(2), 287.

Remland, M. S., Jones, T. S., & Brinkman, H. (1995). Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age [PMID: 7650932]. *The Journal of Social Psychology*, *135*(3), 281.

Rimmele, U., Meier, F., Lange, T., & Born, J. (2010). Suppressing the morning rise in cortisol impairs free recall. *Learning & Memory*, *17*(4), 186.

Romano, J. M., Hsiao, K., Niemeyer, G., Chitta, S., & Kuchenbecker, K. J. (2011). Human-inspired robotic grasp control with tactile sensing. *IEEE Transactions on Robotics*, *27*(6), 1067.

Romano, J. M., & Kuchenbecker, K. J. (2011). Please touch the robot. hands-on demonstration. *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*.

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of personality and social psychology*, *87*(4), 494.

Šabanovic, 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. *Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR)*.

Salimetrics, L., & SalivaBio, L. (2011). Saliva collection and handling advice. *State College, PA: Salimetrics LLC, SalivaBio LLC*, 1.

Salter, T., Davey, N., & Michaud, F. (2014). Designing & developing queball, a robotic device for autism therapy. *The 23rd IEEE international symposium on robot and human interactive communication*, 574.

Salvini, P., Ciaravella, G., Yu, W., Ferri, G., Manzi, A., Mazzolai, B., Laschi, C., Oh, S.-R., & Dario, P. (2010). How safe are service robots in urban environments? bullying a robot. *19th international symposium in robot and human interactive communication*, 1.

Santrock, J., Deater-Deckard, K., & Lansford, J. (2021). *Child development: An introduction. 15th edition*. McGraw Hill.

Sarang, P., & Telles, S. (2006). Effects of two yoga based relaxation techniques on heart rate variability (hrv). *International Journal of Stress Management*, *13*(4), 460.

Scott, T. (2020). 1,204,986 votes decided: What is the best thing? *YouTube*.

Shiomi, M., Nakagawa, K., Shinozawa, K., Matsumura, R., Ishiguro, H., & Hagita, N. (2017). Does a robot's touch encourage human effort? *International Journal of Social Robotics*, 5.

Shiomi, M., Nakata, A., Kanbara, M., & Hagita, N. (2017a). A hug from a robot encourages prosocial behavior. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 418.

Shiomi, M., Nakata, A., Kanbara, M., & Hagita, N. (2017b). A robot that encourages self-disclosure by hug. *Proceedings of the International Conference on Social Robotics (ICSR)*, 324.

Singh, D. (1993). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of personality and social psychology*, *65*(2), 293.

Smarr, C.-A., Mitzner, T. L., Beer, J. M., Prakash, A., Chen, T. L., Kemp, C. C., & Rogers, W. A. (2014). Domestic robots for older adults: Attitudes, preferences, and potential. *International journal of social robotics*, *6*(2), 229.

Statman-Weil, K. (2015). Preschool through grade 3: Creating trauma-sensitive classrooms. *YC Young Children*, *70*(2), 72.

Stiehl, W. D., Lieberman, J., Breazeal, C., Basel, L., Cooper, R., Knight, H., Lalla, L., Maymin, A., & Purchase, S. (2006). The Huggable: A therapeutic robotic companion for relational affective touch. *Proceedings IEEE Consumer Communications and Networking Conference*, 1290.

Stiehl, W. D., Lieberman, J., Breazeal, C., Basel, L., Lalla, L., & Wolf, M. (2005). Design of a therapeutic robotic companion for relational, affective touch. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 408.

Sumioka, H., Nakae, A., Kanai, R., & Ishiguro, H. (2013). Huggable communication medium decreases cortisol levels. *Scientific Reports*, *3*(1), 3034.

Suvilehto, J. T., Glerean, E., Dunbar, R. I. M., Hari, R., & Nummenmaa, L. (2015). Topography of social touching depends on emotional bonds between humans. *Proceedings of the National Academy of Sciences*, *112*(45), 13811.

Suvilehto, J. T., Nummenmaa, L., Harada, T., Dunbar, R. I., Hari, R., Turner, R., Sadato, N., & Kitada, R. (2019). Cross-cultural similarity in relationship-specific social touching. *Proceedings of the Royal Society B*, *286*(1901), 20190467.

Szeto, A., McCabe, P. M., Nation, D. A., Tabak, B. A., Rossetti, M. A., McCullough, M. E., Schneiderman, N., & Mendez, A. J. (2011). Evaluation of enzyme immunoassay and radioimmunoassay methods for the measurement of plasma oxytocin. *Psychosomatic medicine*, *73*(5), 393.

Tackett, S., Young, J. H., Putman, S., Wiener, C., Deruggiero, K., & Bayram, J. D. (2018). Barriers to healthcare among muslim women: A narrative review of the literature. *Women's Studies International Forum*, *69*, 190.

Teh, J. K. S., Cheok, A. D., Peiris, R. L., Choi, Y., Thuong, V., & Lai, S. (2008). Huggy Pajama: A mobile parent and child hugging communication system. *Proceedings of the ACM International Conference on Interaction Design and Children*, 250.

Townsend, W. (2000). The BarrettHand grasper: A programmably flexible part handling and assembly. *Industrial Robot: An International Journal*, *27*(3), 181.

Trafton, J. G., Schultz, A. C., Perznowski, D., Bugajska, M. D., Adams, W., Cassimatis, N. L., & Brock, D. P. (2006). Children and robots learning to play hide and seek. *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 242.

Trost, M. J., Chrysilla, G., Gold, J. I., & Mataric, M. J. (2020). Socially-assistive robots using empathy to reduce pain and distress during peripheral IV placement in children (F. Galli, Ed.) [Publisher: Hindawi]. *Pain Research and Management*, *2020*, 7935215.

Trovato, G., Do, M., Terlemez, Ö., Mandery, C., Ishii, H., Bianchi-Berthouze, N., Asfour, T., & Takanishi, A. (2016). Is hugging a robot weird? investigating the influence of robot appearance on users' perception of hugging. *Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 318.

Tsetserukou, D. (2010). HaptiHug: A novel haptic display for communication of hug over a distance. *Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, 340.

Tuzel, O., Porikli, F., & Meer, P. (2007). Human detection via classification on Riemannian manifolds. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1.

Uvnäs-Moberg, K., Handlin, L., & Petersson, M. (2014). Self-soothing behaviors with particular reference to oxytocin release induced by non-noxious sensory stimulation. *Frontiers in Psychology*, *5*, 1.

VandenBerg, N. L. (2001). The use of a weighted vest to increase on-task behavior in children with attention difficulties. *The American Journal of Occupational Therapy*, *55*(6), 621.

Viola, P., & Jones, M. (2004). Robust real-time object detection. *International Journal of Computer Vision*, *57*(2), 137.

Viola, P., & Jones, M. J. (2001). Robust real-time face detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, *2*, 747.

Vo, D. M., Jiang, L., & Zell, A. (2014). Real time person detection and tracking by mobile robots using RGB-D images. *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 689.

Wallis, M., Popat, S., & McKinney, J. (2010). Embodied conversations: Performance and the design of a robotics dancing partner. *Design Studies*, *31*(2), 99.

Wang, Z., Peer, A., & Buss, M. (2009). An HMM approach to realistic haptic human-robot interaction. *Proceedings of the IEEE World Haptics Conference (WHC)*, 374.

Watts, C., & Zimmerman, C. (2002). Violence against women: Global scope and magnitude. *The Lancet*, *359*(9313), 1232.

Weiss, A., Bernhaupt, R., Tscheligi, M., Wollherr, D., Kuhnlenz, K., & Buss, M. (2008). A methodological variation for acceptance evaluation of human-robot interaction in public places. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 713.

Wilhelm, I., Born, J., Kudielka, B. M., Schlotz, W., & Wüst, S. (2007). Is the cortisol awakening rise a response to awakening? *Psychoneuroendocrinology*, *32*(4), 358.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322.

Wirobski, G., Schaebs, F., Range, F., Marshall-Pescini, S., & Deschner, T. (2021). Analytical and physiological validation of an enzyme immunoassay to measure oxytocin in dog, wolf, and human urine samples. *Scientific Reports*, *11*(1), 1.

Yamane, K., Kim, J., & Alspach, A. N. (2017). Soft body robot for physical interaction.

Yamazaki, R., Christensen, L., Skov, K., Chang, C.-C., Damholdt, M. F., Sumioka, H., Nishio, S., & Ishiguro, H. (2016). Intimacy in phone conversations: Anxiety reduction for Danish seniors with Hugvie. *Frontiers in Psychology*, *7*, 537.

Yaniv, A. U., Salomon, R., Waidergoren, S., Shimon-Raz, O., Djalovski, A., & Feldman, R. (2021). Synchronous caregiving from birth to adulthood tunes humans' social brain. *Proceedings of the National Academy of Sciences*, *118*(14).

Yohanan, S., & MacLean, K. E. (2012). The role of affective touch in human-robot interaction: Human intent and expectations in touching the Haptic Creature. *International Journal of Social Robotics*, *4*(2), 163.

Yohanan, S. J. (2012). *The Haptic Creature: Social human-robot interaction through affective touch* (Doctoral dissertation). University of British Columbia.

Yoshida, S., Kawahara, Y., Sasatani, T., Kiyono, K., Kobayashi, Y., & Funato, H. (2020). Infants show physiological responses specific to parental hugs. *iScience*, *23*(4), 100996.

Yurtsever, E., Lambert, J., Carballo, A., & Takeda, K. (2020). A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access*, *8*, 58443.