TRUST ATTRIBUTION IN COLLABORATIVE ROBOTS: AN EXPERIMENTAL INVESTIGATION OF NON-VERBAL CUES IN A VIRTUAL HUMAN-ROBOT INTERACTION SETTING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AHMET MERIÇ ÖZCAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COGNITIVE SCIENCE

JUNE 2021

TRUST ATTRIBUTION IN COLLABORATIVE ROBOTS: AN EXPERIMENTAL INVESTIGATION OF NON-VERBAL CUES IN A VIRTUAL HUMAN-ROBOT INTERACTION SETTING

submitted by **AHMET MERIÇ ÖZCAN** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science Department, Middle East Technical University** by,

Prof. Dr. Deniz Zeyrek Bozşahin
Dean, **Graduate School of Informatics** ________________

Dr. Ceyhan Temürcü
Head of Department, **Cognitive Science** ________________

Assoc. Prof. Dr. Erol Şahin
Supervisor, **Computer Engineering Dept., METU** ________________

Assoc. Prof. Dr. Cengiz Acartürk
Co-supervisor, **Cognitive Science Dept, METU** ________________

Examining Committee Members:

Assist. Prof. Dr. Barbaros Yet
Cognitive Science Dept., METU ________________

Assoc. Prof. Dr. Erol Şahin
Computer Engineering Dept., METU ________________

Assoc. Prof. Dr. Cengiz Acartürk
Cognitive Science Dept., METU ________________

Assist. Prof. Dr. Murat Perit Çakır
Cognitive Science Dept., METU ________________

Assoc. Prof. Dr. Emre Uğur
Computer Engineering Dept., Boğaziçi University ________________

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Ahmet Meriç Özcan

Signature :

ABSTRACT

TRUST ATTRIBUTION IN COLLABORATIVE ROBOTS: AN EXPERIMENTAL INVESTIGATION OF NON-VERBAL CUES IN A VIRTUAL HUMAN-ROBOT INTERACTION SETTING

Özcan, Ahmet Meriç
M.S., Department of Cognitive Science
Supervisor: Assoc. Prof. Dr. Erol Şahin
Co-Supervisor: Assoc. Prof. Dr. Cengiz Acartürk

June 2021, 61 pages

This thesis reports the development of non-verbal HRI (Human-Robot Interaction) behaviors on a robotic manipulator, evaluating the role of trust in collaborative assembly tasks. Towards this end, we developed four non-verbal HRI behaviors, namely gazing, head nodding, tilting, and shaking, on a UR5 robotic manipulator. We used them under different degrees of trust of the user to the robot actions. Specifically, we used a certain "head-on neck posture" for the cobot using the last three links along with the gripper. The gaze behavior directed the gripper towards the desired point in space, alongside with the head nodding and shaking behaviors. We designed a remote setup to experiment subjects interacting with the cobot remotely via Zoom teleconferencing. In a simple collaborative scenario, the efficacy of these behaviors was assessed in terms of their impact on the formation of trust between the robot and the user and task performance. 19 people participated in the experiment with varying ages and genders.

Keywords: robotics, human-robot interaction, non-verbal gestures, telepresence robot

ÖZ

KOLABORATİF ROBOTLARDA GÜVEN ÖZELLİĞİ: SANAL İNSAN ROBOT ETKİLEŞİM ORTAMINDA, SÖZSÜZ İPUÇLARININ DENEYSEL ARAŞTIRMASI

Özcan, Ahmet Meriç
Yüksek Lisans, Bölümü
Tez Yöneticisi: Doç. Dr. Erol Şahin
Ortak Tez Yöneticisi: Doç. Dr. Cengiz Acartürk

Haziran 2021 , 61 sayfa

Bu tez insan robot arası etkileşimi geliştirmek amacıyla, yardımcı UR5 robotunun manipülatörü ile, bakış ve kafa davranışları yaratmayı ve etkilerini montaj senaryosu altında test etmeyi hedeflemektedir. Bu doğrultuda çeşitli sözlü olmayan robot davranışları UR5 robotu ve Robotiq çene kıskacı kullanılarak geliştirildi, bunlar; yana ve öne kafa sallama, kafa eğme ve bakış davranışıdır. Bu davranışları uygulayabilmek için daha önceden dizayn edilmiş bir robot duruşu kullanıldı ve son üç robot eklemi, çene kıskacı kullanılarak "baş-boyun" yapısına çevrildi. Bu duruş yapısı ile birlikte çene kıskacı uzayda bir noktaya doğrultularak bakış davranışı yapabilmektedir. Bakış davranışına ek olarak kafa yapısı ile birlikte kafa sallama gibi davranışlarda modellendi, bunun yanında katılımcıların aktif olarak cobot ile birlikte telekonferans programı olan Zoom üzerinden etkileşime geçebileceği özgün bir deney ortamı geliştirildi. Ortak çalışmaya dayalı bir senaryoda bu davranışların güven kazanımı ve performans üzerindeki etkisi test edildi. Farklı yaş ve cinsiyet gruplarından 19 katılımcı ile birlikte deneyler gerçekleştirildi.

Anahtar Kelimeler: robotik, insan robot etkileşimi, uzak bağlantı robot deneyi

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to Erol Sahin and Cengiz Acarturk, for supervising my M.Sc. studies. Their patience, enthusiasm, extensive visions and deep knowledge have encouraged me through my academic research.

My appreciation also goes out to my family and friends for their encouragement and support all through my studies.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

HRI	Human robot interaction
HIH	Human-to-Human
AHI	Animal Human Interaction
DOF	Degrees of freedom
GIL	Global Interpreter lock
GUI	Graphical User Interface
UDP	User Datagram Protocol
ROS	Robotic Operating System
IPC	Inter Process Communication

CHAPTER 1

INTRODUCTION

The concept of trust has been studied in various domains of human-robot interaction, including autonomous vehicles and social robots for rehabilitation (Langer, Feingold-Polak, Mueller, Kellmeyer, & Levy-Tzedek, 2019). In social robotics, establishing a trustworthy interaction has a significant role in effective communication between a robot and its users. At a broader level, trust is a cognitively complex concept that is difficult to measure and implement. It is insufficient to ask people where they place their trust (Botsman, 2017). For a systematic investigation of trust in human-robot interaction, we need quantifiable measures of trust, such as the duration of task completion or the accuracy of a task under various experimental manipulation of trust conditions. The goal of the present study is to study trust within the context of the gestures produced by a collaborative robotic manipulator, as measured in terms of task performance in a collaborative task. Accordingly, we expect to find a measurable impact of the degree of trustworthiness of the robot on human task performance.

Collaborative robotic manipulators, a.k.a. cobots, are designed to carry out manipulation tasks directly interacting with workers as their work-mates (Colgate, Wannasuphoprasit, & Peshkin, 1996). This requires cobots to establish and maintain short- and long-term interactions with their human co-workers through social cues. However, current cobots are designed as simple robotic manipulators equipped with safety features and do not provide any explicit support for human-robot interactions.

Human-robot interaction is the study of investigating, evaluating, and designing robotic systems for humans. Such systems require natural communication (a way of most people exchange information, e.g. speech and gestures) (Admoni & Scassellati, 2017), including verbal (writing, speaking, etc.) and non-verbal (gestures, appearance, posture, gaze etc.) techniques. In our thesis we focused on several commonly used non-verbal interaction techniques (Cha, Kim, Fong, & Mataric, 2018)(Admoni & Scassellati, 2017) in HRI: gaze and head gestures.

Gaze behavior is a significant non-verbal communication method. It is regarded as an important cue in social interactions (Cook, 1977), and cognitively influential since there are special "hard-wired" neurological structures for these visual actions. Prior work indicated that even pointing the gripper of a cobot toward its user or toward an object can be perceived as a gaze and facilitate social perceptions and interaction (Terzioğlu, Mutlu, & Şahin, 2020). Therefore we also embraced this non-verbal cue and used it in our experiments. However, addition to prior work, we tried to mathematically model this behavior for a cobot.

Along with the gaze, head gestures are also another important non-verbal communication method, that is commonly used in HRI (Sidner, Lee, Morency, & Forlines, 2006) (Liu, T., et al., 2012) (Breazeal et al., 2005). In general head gestures convey useful information about the intentions and feelings, and have a significant place in one-to-one social communication (Morency, Sidner, Lee, & Darrell, 2007). For example, head nodding behavior is interpreted as agreement or approval and head shake behavior is denial or disagreement in many cultures even though they are not universally employed (Darwin, n.d.).

In previous work (Terzioğlu et al., 2020), it has been studied how the HRI quality between the human worker and a cobot can be enhanced by applying some of the animation principles of Disney(Lasseter, 1987), such as appeal, secondary behavior, and arc'ing. Towards increasing the appeal of the cobot, a "head-on-neck posture" for the cobot, affixed a sunglass on the gripper, and pointed it towards the human or a common point in space to generate a gaze, as shown in Figure 1.1b.

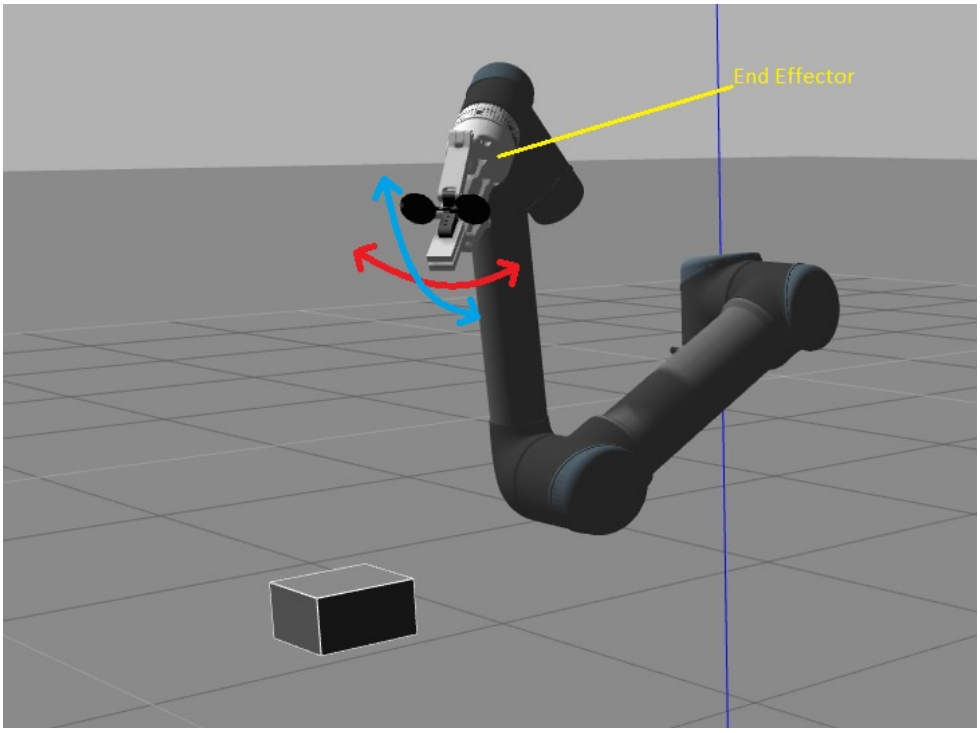
By taking prior work to account, in this thesis, we implemented four non-verbal behaviors on "head-on-neck posture", on the cobot platform; namely *gaze*, *head nod* *head tilt* and *head shake* and investigated how gaze and head shakes with gaze aversion, would affect the trust attribution in cobots in a virtual human-robot interaction setting. The formation of trust has been an intense topic of research in various disciplines, including social sciences and more recently Artificial Intelligence under different connotations of the concept. In the present study, we investigate the trust concept in terms of the trustworthiness of the robot's gaze-pointing actions and we measured the task performance of the participants.

1.1 Terminology

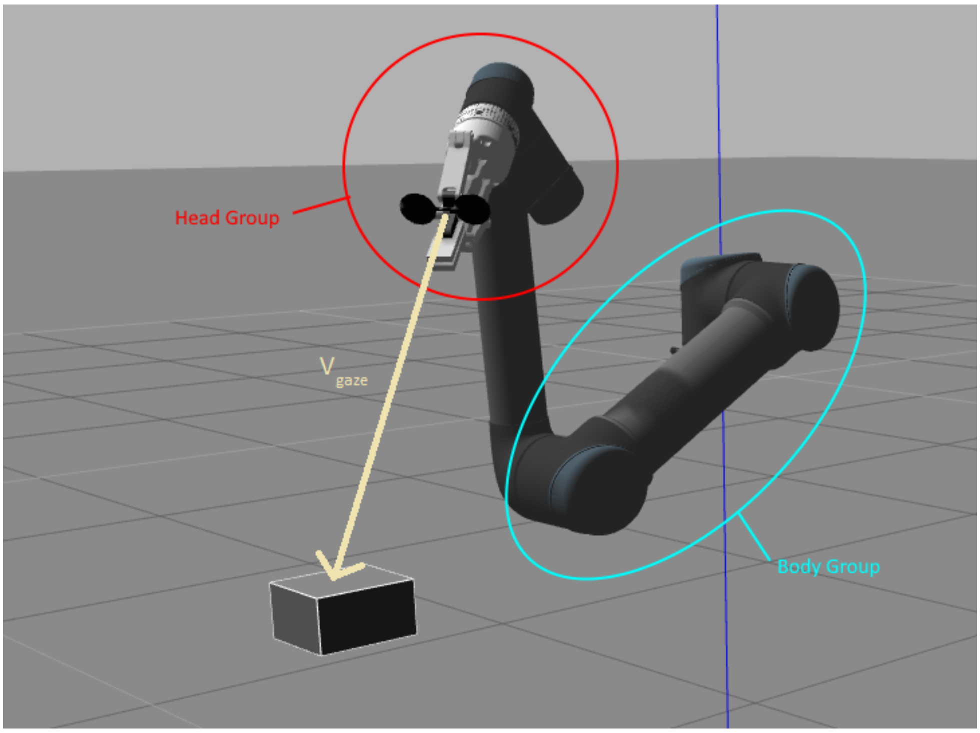
In this section, the terminology used throughout the thesis is defined.

- **Collaborative Universal Robot**s are commercial robots that provide a cost-effective, flexible, and safe automation solution for a wide range of production tasks and UR5 Collaborative robot is one of the members of these robots. UR5 robot is shown in the Figure 1.1a.
- **Assembly Task** is defined as the process of follows a pre-defined sequence of assembling actions towards assembling a product.
- **Gaze** is defined as the steady look to a certain point in space, which is assumed to indicate the attention of the human/robot.
- **Gaze Vector** is a vector that starts at the center of the eyes and ends at the point to look.
- **End-effector** is a part of the robot which manipulates and interacts with the outside world.
- **Body Group Joints** are the first three joints of the UR5 Robot, which acts as a UR5 robot's body.

- **Head Group Joints** are the last three joint of UR5 Robot, which acts as head of UR5 Robot.
- **Head Shake** is the horizontal motion of the head around the gazing direction, usually conveys a negative response to an action in most cultures. The movement illustrated in Figure 1.1a with red arrows.
- **Head Nod** is the vertical motion of the head around the gazing direction, is commonly used for conveying a positive response. The movement illustrated in Figure 1.1a with blue arrows.
- **Worker** is a human participant who involves an assembly process.
- **Handover** is a complicated collaborative task, "where actors coordinate in time and space to transfer control of an object." (Strabala et al., 2013)



(a) Vertical and horizontal movement of End-Effector



(b) Gaze vector, Head group and Body Group illustration

Figure 1.1: UR5 Collaborative robot definitions

1.2 Research Questions

The major experimental research question of this thesis is to understand the relationship between level of trust worthiness, and performance of the participants on an assembly task. We hypothesise that the level of trust have an impact on collaborative task where robots and human working together to produce piece of work. Through this aim, gaze, gaze aversion and head shaking behaviours choosen from our developed gestures and used inside a novel interactive experiment setup environment over zoom teleconference (*Video Conferencing, Web Conferencing, Webinars, Screen Sharing*, n.d.). This setup allowed people to interact with a cobot over the internet to participate a HRI experiment. A virtual table with several holes were presented to the participants along with the screws to drive inside the holes. The task of the participant was to drive as many screws as possible into the correct holes in given time. UR5 collaborative robot gave gaze cues to the participants thought out the experiment with different trust factors, %0 %50 and %100 as an independent variable. We investigated the human participants performance through the virtual experiment trials, for average screw time, total number of correct and miss-placed screws. In our hypothesis, we believe that the robot's cues will improve the task performance for the quantitative measures depending on the trust rates of the human participants. In other words we think that the task performance will have a correlation with the robot's trust worthiness.

CHAPTER 2

LITERATURE REVIEW

In this chapter, we present the relevant literature for gaze and other head gestures in HRI.

2.1 Gaze in HRI

Human eyes are considered to be the windows of our mind hence gazing directly from our eyes on some point in space is a very informative signal. Consequently gaze is substantial non-verbal communication tool for humans. Biological findings assert that, along with humans, many vertebrate species can comprehend gaze (Emery, 2000). Humans, can comprehend intentions' of others much more successfully than vertebrate species (Emery, 2000). Two centuries ago Darwin stated the significance of gaze on understanding others emotions and feelings by arguing that gaze can reveal, pride, humility, guilt suspicion etc. (Darwin, n.d.). It is also essential behavior for social communications. It gives several different cues by revealing target point of attention (people look at the person that they listening in a conversation, %88 percent of the time (Vertegaal, Slagter, Veer, & Nijholt, 2001) (Cook, 1977) and turn taking in a conversation (Rossano, 2012).

Because of the significance of gaze, it has been used under several different research topics. The precursor researches about gaze was initiated by the virtual agent community, which uses computer programs to animate agents, in early 1990s (Admoni & Scassellati, 2017). Virtual agents with gaze capability has shown improvement on means of capturing attention, maintaining engagement, and contributing conversational fluidity along humans (Cassell, 2000 April).

Robotics began introducing meaningful eye gaze into their systems in the late 1990s, with robots; namely Cog (Scassellati, 1996) Kısmet (Breazeal & Scassellati, n.d.). Modern-day approaches on gaze diversify, and it has been used for several different topics such as artificial intelligence, virtual agents, psychology (Admoni & Scassellati, 2017).

Throughout our research we came across various types of gaze behaviours:

- **Joint Gaze (Attention)** behaviour involves sharing common gaze to a certain point or an object. (Moore & Dunham, 2014). For an example: in assembly task it occurs when a robot and a human worker gaze a piece or a tool together.

- **Mutual Gaze (Attention)** also refer to "eye contact", which involves mutually looking to each others eyes.

- **Gaze aversion** is changing the main direction of gaze while facing directly to an agent's face. It is an important nonverbal cue that serves number of interaction functionalities, adjusting intimacy, managing conversational awareness (Andrist, Tan, Gleicher, & Mutlu, 2014).

In the coming subsections, we will discuss common topics found during our research on gaze in HRI. By identifying the commonalities, these subsections will emphasize varying appearances, the embodiment of agents, and functionalities of gaze behavior.

2.1.1 Functionalities of Gaze Behaviour

Gaze behavior is used for several different purposes: collecting information, indicating emotional state and interest, directing conversations (e.g., through turn-taking), shifting or showing attention, establishing trust link and more (Ruhland et al., 2015). In a certain context and with necessary abilities (e.g., moving head or gaze for to a specific location to shift attention), like humans, robots and virtual agents can also be able to perform these functionalities (Ruhland et al., 2015).

2.1.1.1 Gaze on Attention

Gaze is a commonly exploited behavior to shift and direct the attention of others (Admoni & Scassellati, 2017). One way of doing this is gaze cuing. Gaze cue is the behavior in which we pay attention to someone else's gaze and find our gaze drawn by it. In HRI, gaze cue behavior can be achieved by moving head orientation (by moving gaze together with orientation) or eyeballs (Hoque et al., 2013). It is a powerful method that shifts the attention of others even if they are not facing directly to an agent (Hoque et al., 2013).

Figure 2.1: Humanoid robot face (capable of moving eyes), which is used for catching attention of users (Hoque et al., 2013)

Gaze cues were also established with NAO robots (*NAO the humanoid and programmable robot: SoftBank Robotics*, n.d.) which is incapable of moving eyes (Mwangi et al., 2018). In order to reallocate their gaze, NAO robots use their head movements. The experiment used gaze cues in the context of a shell game in which an object is hidden under one of three cups, and those cups are shuffled at a certain speed depending on different difficulty levels. It was asked participants to find the cup where the object is hidden. Human participants were able to capture NAO robots' gaze cues and improved their response time. In figure 2.2 the setup environment is illustrated.

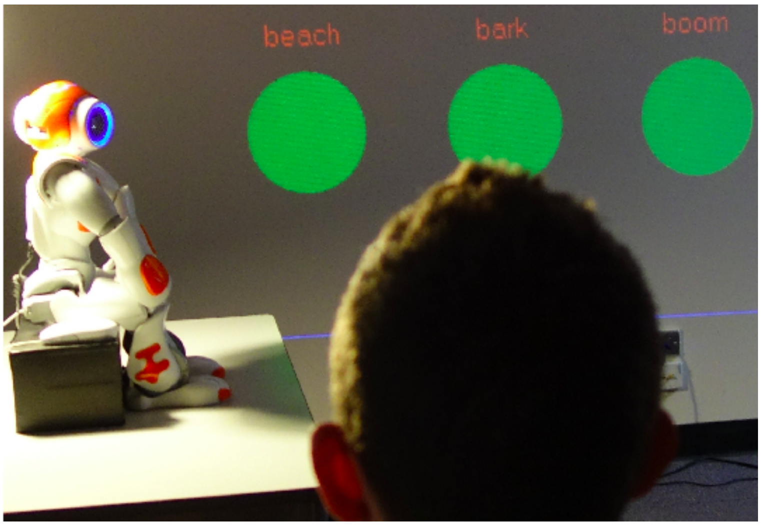


Figure 2.2: NAO Robots shell game experiment environment (Stanton & Stevens, 2014)

Similar to the previous experiment, NAO robots' gaze capabilities were tested in another game, pair card game. Fourteen cards were arranged and closed down on a table between two players as in figure 2.4. Participants' task was defined as finding the pair cards as fast as possible, and NAO robots tried to aid participants with their gaze cues. Participants' eye movements were tracked and observed through eye trackers. This demonstrated that NAO robots and human gaze could achieve similar time duration for task solving (Mwangi et al., 2018).

Figure 2.3: NAO robots Pairing Card Game (Mwangi et al., 2018)

Along being the primary source of information channel, gaze cues can also enhance explicit communication techniques (Breazeal et al., 2005) such as pointing an object with a hand. Explicit communication is defined as deliberate where a person aims to share specific information. On the other hand, implicit behaviors are defined as transmitting information that inherits in a gesture such as gaze. The Humanoid robot Leonardo (in Figure 2.4) was used in an experiment to understand the effects of implicit behaviors where he used his explicit + implicit (e.g. Gaze cue) and explicit gestures alone, in two different setups. Experiment results indicated that participants were confused when the robot pointed an object without an implicit gaze behavior (Breazeal et al., 2005).

Figure 2.4: Leonardo Robot Gaze (Breazeal et al., 2005)

2.1.1.2 Gaze on Emotions and Expressions

A robot could benefit from the ability to express personality or emotion. For example, robots that interact with humans for a long duration of time should have a pleasant personality, or helping robots should be trustworthy for their human companions.

Gaze is one of the many non-verbal behaviors that can express the personalities and emotions of the user. Occurrence of gaze types inside a certain context can create engaging interaction or mean negative emotions, such as gaze duration can generate intimacy embarrassment, self-disclosure, or attraction (Kleinke, 1986). As Mehrabian stated, people tend to gaze more at the people who they liked than disliked, which they explained it as intimacy of gaze (Mehrabian, 1968). People also stare less with interviewers when they are asked personal and embarrassing questions (Exline, Gray, & Schuette, 1965).

In HRI, similar to human-human experiments gaze aversion can show feeling of distrust (Normoyle et al., 2013) and withdrawn personalities (Andrist, Mutlu, & Tapus, 2015) or it can be utilized to seen as a thoughtful and creative agent (NAO robots were used in an experiment, where they avert gaze before answering to participants' questions) (Andrist et al., 2014), depending on the context. Along with these expressions and emotions, virtual agents' eye movements could be able interpreted as different types of feelings, such as joy, fear, anger, disgust, and surprise (Z. Li & Mao, 2012).

2.1.2 Embodiment and Virtual Agents on Gaze

Precursor research with virtual agents led the way for embodied gaze research on robotics (Admoni & Scassellati, 2017) and improved the understanding of gaze in HRI (Ruhland et al., 2015).

The affect of gaze can be examined under three different category of agents: namely, copresent (physically embodied as well as physically present in a user's environment(J. Li, 2015), telepresent (physically embodied but displayed on a computer screen (J. Li, 2015) and virtual agents. It was shown that that physically copresent embodied systems improve interactions over virtual and telepresent systems (J. Li, 2015). However it was also emphasized that there was no significant difference between telepresent and virtual agents (J. Li, 2015).

On the other hand, compared to robots, virtual agents can better control their bodies' movement and gaze timing. Virtual agents are easily programmed to control their facial expressions, mimics, eyebrows, eyelids, etc. Furthermore virtual agents could add different body gestures along with the gaze behaviour (Ruhland et al., 2015).

2.1.3 Robot Appearance on Gaze

Gaze experiments conducted using robots with a range of diversity on appearance and abilities. These setups vary non-humanoid robots to extremely life-like humanoid robots and agents (Breazeal & Scassellati, n.d.) (Zaraki et al., 2014) (Szafir & Mutlu, 2012). Additionally, robots with varying gaze capabilities are also used. For example, some robots can move their eyes balls, some have virtual faces through a computer screen (Fitter, Mohan, Kuchenbecker, & Johnson, 2020), and some could blink (Breazeal & Scassellati, n.d.) etc.

Varying gaze capabilities depends on the to the high cost of producing eye movements on robots. Every single moving part of the robot must be generated by some motor or actuator and these actuator must be small to fit in robot and powerful to make rapid movements. In other words, gaze capabilities, adjusts robot's complexity and cost, thus people bear these requirements in mind to design their robots.

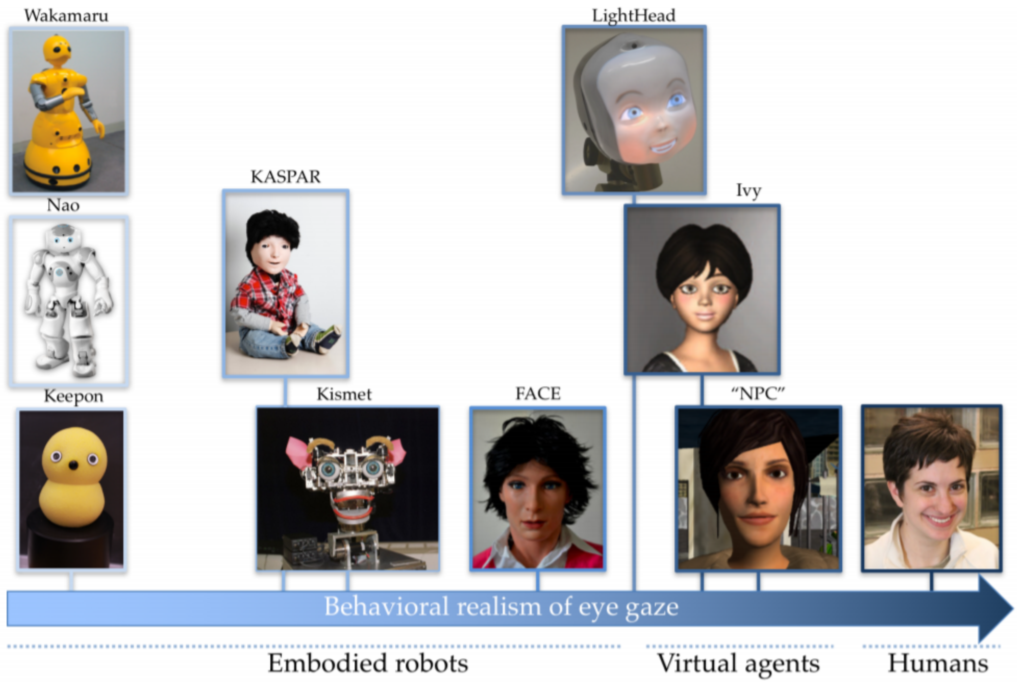


Figure 2.5: The Spectrum of realism factor of a robot with varying appearance and capabilities. Robots in the figure are: Wakamaru (Szafir & Mutlu, 2012), NAO Robot[1], Keepon KASPAR (courtesy of the Adaptive Systems Research Group, University of Hertfordshire, UK), Kismet (Breazeal & Scassellati, n.d.), FACE (Zaraki et al., 2014), LightHead (Delaunay & Belpaeme, 2012), Ivy (Andrist et al., 2013), and an NPC (Normoyle et al., 2013). (The image taken from (Admoni & Scassellati, 2017))

[1] Nao Robots designed by the company Aldebaran which is now Soft Bank Robotics https://www.softbankrobotics.com/

The appearance and capability create a realism factor for face, and eye gaze (Admoni & Scassellati, 2017). When the complexity increases and the more they resemble a living animal or a human, the robot's eye gaze becomes more realistic. Figure 3.1 illustrates a robot's realism factor. The most right end side of the realism factor spectrum comprises an actual living human, and when you go just below a human figure, there are some cartoonish virtual agents. Further, you go on the left side of the spectrum; there are other robots with face gestures and eye movement capabilities such as Kismet; however, they are starting to lose complexity and face capabilities. In general, robots on the right side can transfer more information with their faces since they have more facial features. Robots on the extreme left side of the spectrum generally do not have these features, such as Keepon, NAO robot, and Wakamuru. However, their faces resemble animals or humans, and they can also generate gaze stimuli. This logic is similar to the pareidolia phenomenon. Pareidolia phenomenon suggests that humans tend to find patterns, face like in our case, known for them. Thus, even with these robots, which have static eyes and simplistic looks, we can perceive their gaze.

Robot appearance directly impacts several different variables such as likeability, sociability, perceived intelligence, etc. Not every robot is considered socially advanced or likable, and also, some robots causing uncanny valley as Mori defined (Mori, 1993). Uncanny valley is a certain area in which robots look both human-like and robot-like simultaneously. This region becomes confusing for humans to categorize agents. Hence Mori argued that humans tend to prefer a human-like over a robot-like appearance. Additionally, Mori also stated that robot movements steepens the slope of the uncanny valley (in figure 2.6) (Mori, 1993). UR5 Collaborative robot is located on the far left side of the affinity graph, and properly designed movements improve its affinity.

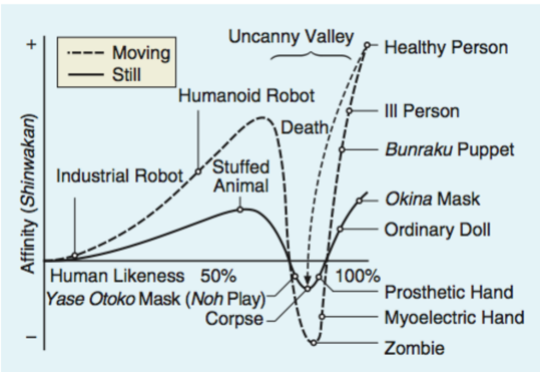


Figure 2.6: The Affect of movement on uncanny valley (Mori, 1993)

2.2 Head Gestures in HRI

Head gestures are a popular form of nonverbal communication method defined by a certain continuous movement of the head in a time period. They are used across various interactions such as conversation, collaborative tasks, instructions, etc. Meaning in these interactions might change since they are evolved differently across cultures

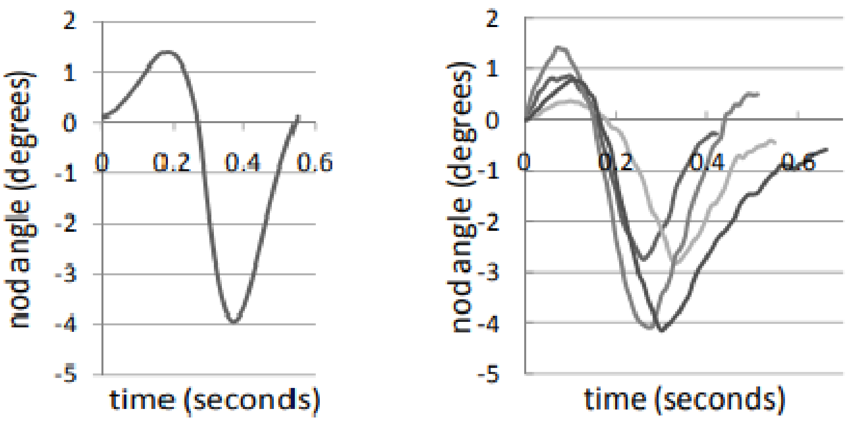
(Darwin, n.d.). Therefore head gestures generally vary not only in type but context and social significance. Even though they convey semantic information, they can enhance narration, regulate conversations, emphasize what's being said (Cha et al., 2018)

In this review we mainly focused on several head gestures: namely head nodding, head shake, and head tilt.

2.2.1 Head Nodding

Head nodding is a common non-verbal behavior in different cultures, meaning agreement or understanding (Darwin, n.d.). There are several contrasting ideas about the roots of nodding; as Darwin stated, head nodding behavior also originated from infants when they try to get food into their mouths. They often inclined their heads forward (Darwin, n.d.) hence this action resembled the head nodding behavior. However, several types of research about head nodding argued the opposite, as blind children do not use head nodding behavior (Iverson, Tencer, Lany, & Goldin-Meadow, 2000) hence it might be learned behavior. To this day, the origin of head nodding behavior continues to puzzle researchers.

In HRI, head nodding behavior is also used on varying types of robots. (Liu, Ishi, et al., 2012)(Breazeal et al., 2005) (Sidner et al., 2006) To generate a natural head nodding behaviour on robots, a time and angle model is needed for the movement of the head. Figure 2.7b represents several different samples from humans head nodding behaviour from a database (Liu, Ishi, et al., 2012). The duration of the nodding varies between 0.4 to 0.7 seconds, and firstly there is a slight upward motion which generally occurs before the down-up motion. Additionally, the maximum angle reached by upward motion is slightly less than the down-up motion. In figure 2.7a, the dataset samples averaged into one model for head nodding behavior.



(a) Average Head Nod

(b) Head Nodding behaviour samples

Figure 2.7: Head Nodding behaviour (Liu, Ishi, et al., 2012)

2.2.2 Head Shake

A Head shake is a common gesture used in many cultures and implies refusal, denial, and disagreement. It is also a crucial non-verbal behavior that is learned in the early development of an infant (learned at 10.29 months old) even before head nodding (learned at 14.50 months old)("Consistency in Early Vocabulary", 2021). Darwin was among the first close researcher of the gesture, and he proposed, this gesture is originated from how children refuse mothers' nipples, moving their heads side-to-side.

In HRI, this gesture is used for several purposes, such as giving nonverbal cues in a collaborative task. The robot, Leo, used these gestures as an explicit non-verbal cue to the participants when they asked yes/no questions. (Breazeal et al., 2005). Additionally, this behavior is used for several attention shifting experiments with humanoid faces (Hoque et al., 2013).

2.2.3 Head Tilt

Head tilt is an animal and human gesture which is generally done unconsciously (e.g, while posing photographs (Costa & Bitti, 2000)). This behavior also named head canting and head cocking in the literature. It is defined as tilting the head toward one side so that the plane going from the middle of the forehead to mouth is not perpendicular to the horizontal line connecting shoulders (Halberstadt & Saitta, 1987). In figure 2.8 a head tilt behavior illustrated on a human and a dog.

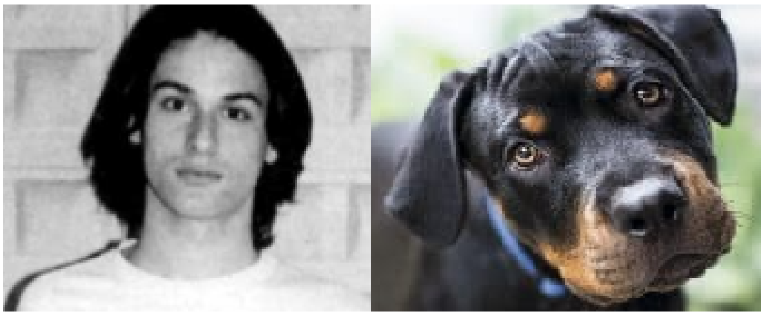


Figure 2.8: Head tilt behaviour on dogs and humans (CarnesMS, 2019) (Costa & Bitti, 2000)

There is no universal reasoning behind head tilt behaviour and it can have various interpretations (Mara & Appel, 2015) (Heads, Llera, & Llera, n.d.). Dogs do this behaviour to understand their surroundings and hear better (by opening their ear flaps) and sometimes it can indicate medical problems (Heads et al., n.d.)(CarnesMS, 2019). People tilt their head in natural interactions settings %40 of the time unconsciously (Halberstadt & Saitta, 1987). However findings shows that this behaviour increase perceived physical attractiveness (Otta, Lira, Delevati, Cesar, & Pires, 1994) (Costa & Bitti, 2000) on humans. In HRI, paralel with the human experiments, the effects of head tilt behaviour is observed (in figure 2.9). It was stated that the influence of head tilt varied depending on the robot. However in general tilt behaviour improved several parameters such as, likeability, cuteness(in figure 2.10) human-likeness etc.(Mara & Appel, 2015)

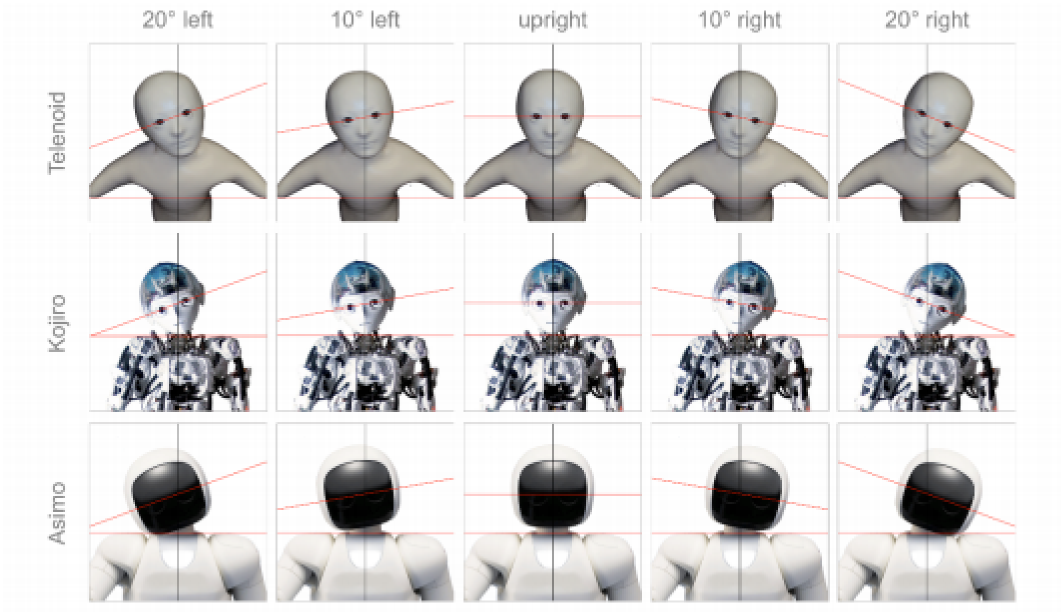


Figure 2.9: Head tilt generation on different robots (Mara & Appel, 2015).

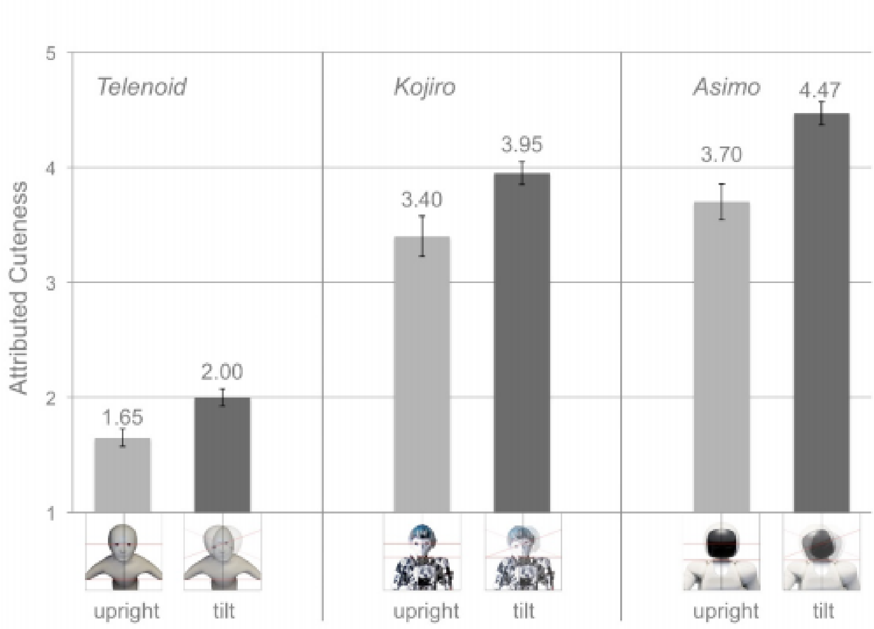


Figure 2.10: The Effect of Head Tilt on different robots. (Mara & Appel, 2015).

CHAPTER 3

METHODOLOGY

In this chapter, the proposed behavioral model for assembly task and mathematical descriptions of the four head gestures will be explained.

3.1 Mathematical Description of Head gestures on UR5

UR5 Collaborative Robot (*Collaborative robotic automation: Cobots from Universal Robots*, n.d.) arm has six Degrees of Freedom (DOF), and it has no natural facial structure. In order to generate head gestures, firstly, a face or head structure should be formed on UR5 robot (Terzioğlu et al., 2020). Additionally, we separated the 6 DOF system into two groups: body and head, to create a "head-on-neck" posture. The body group consists of the first three joints, and the head group, the last three joints. For the head group, black sunglasses was fixated on the manipulator to improve gazing (Terzioğlu et al., 2020). Lastly, for head gestures and gaze behavior, the body group was fixated by constraints, and head group joints were manipulated through some inverse kinematics chains.

3.1.1 Gaze Behaviour

After fixating the body group, an inverse kinematic model is required for the head group orientation using gaze vector. Gaze vector V_{gaze} is defined as the vector beginning at the eyes and ends at the target position in 3D space. In our case the previously attached sunglasses was used as a point of observation for the UR5 collaborative robot. However, the sunglasses' exact position cannot be solved by inverse kinematic equations since it is not part of the UR5 robot inverse kinematic chain. To fix this issue, we assumed that sunglasses' location is P_{j6} (Position of 6^{th} joint) and the target position was defined as P_{target}. In the end, the gaze vector equation is illustrated below.

$$V_{gaze} = \begin{bmatrix} P_{target_x} - P_{j6_x} \\ P_{target_y} - P_{j6_y} \\ P_{target_z} - P_{j6_z} \end{bmatrix} \tag{3.1}$$

After computing the gaze vector, which is defined as V_{gaze}, an inverse kinematic solution is required to align the head group look at a certain position. In an ideal scenario the z-axis of the 6^{th} joint should intersect with the gaze vector, to imitate gaze behavior. However, in our case, it is a little bit more complex because of the UR5 robot's head group joint structure. UR5 Head group joint frames are illustrated in Figure 3.1.

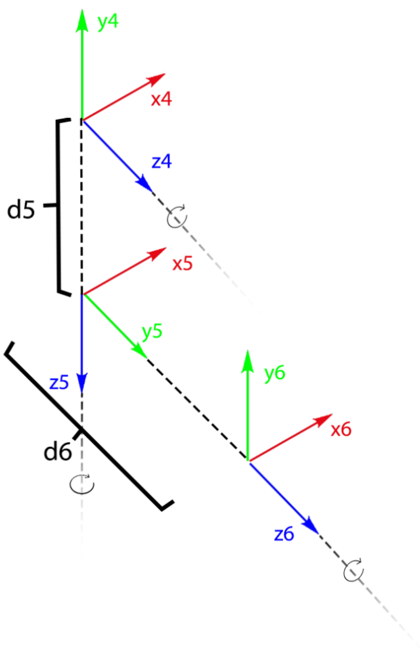


Figure 3.1: Collabrative UR5 robot last three joint frames

In UR5 the last three joints have a non-spherical servo configuration. By definition, a spherical wrist consists of three revolute joints where all of its revolute joints Z-Axes (Z_4, Z_5, Z_6) intersect on the center of the second a revolute joint as it is illustrated in Figure 3.2. This special circumstance creates a sphere equation centered around the second revolute joint. In other words, this allows 3^{th} joint of the spherical wrist to reach every single point on that sphere.This special configuration would let our robot's head group look at every single position by aligning the z-axis of its last revolute joint with the gaze vector.In other words inverse kinematics of a rotation matrix for 3 joint would give the desired behaviour.

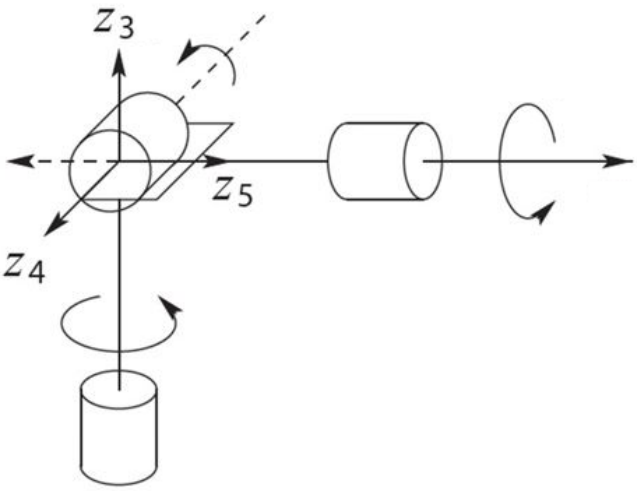


Figure 3.2: Spherical Wrist Configuration

Even though the UR5 Collaborative robot does not have spherical wrist configuration, it is assumed to have one to generate gaze behaviour. As previously explained gaze vector starts from the head group's third joint position (P_{j6}) and ends at the position to look (P_{target}). A gaze behavior is created by aligning the z-axis of the head group's third joint with the gaze vector. In Figure 3.3 the difference between these two vectors identified with a degree, θ. In order to represent this mathematically a rotation matrix, R_{gaze}, is needed and to generate gaze behaviour an inverse kinematic solution is used to compute the rotation matrix.

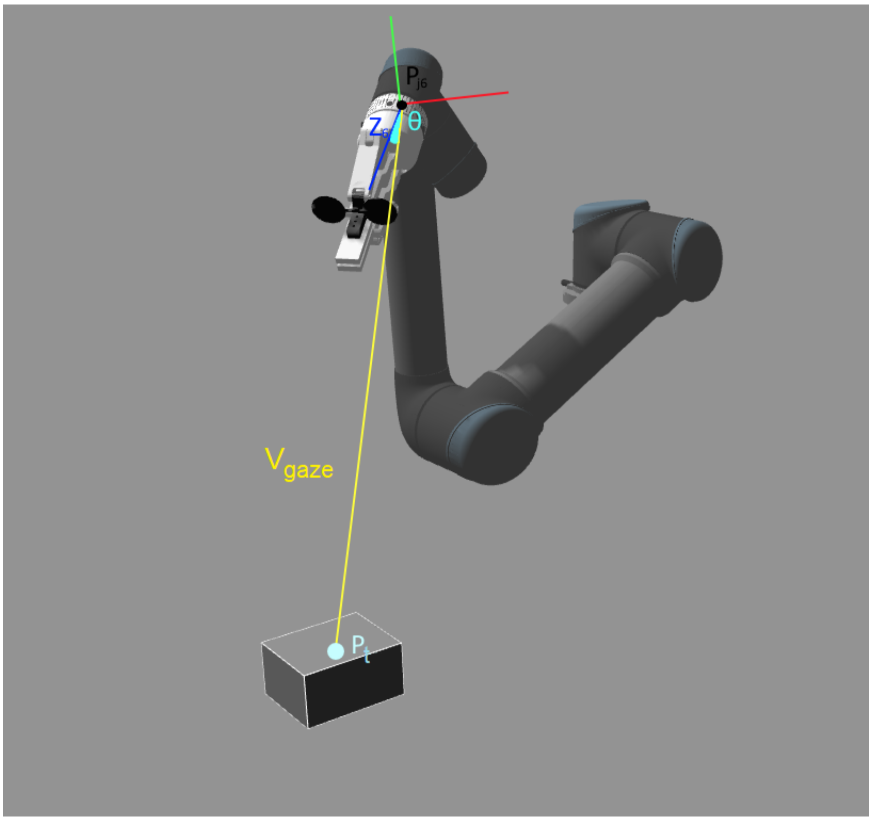


Figure 3.3: The Gaze vector representation on UR5 Collaborative robot

Two different rotation operation is needed to align the last revolute joint's z-axis with the gaze vector. One is on the x-axis, and the other is on the y-axis. By rotating in x and y axis, the orientation of the z-axis is preserved. The rotation matrix, R_{gaze}, depends on, θ_1, rotation in the x-axis and θ_2, rotation in y-axis, which is defined as below:

$$R_{\theta_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{bmatrix}, R_{\theta_2} = \begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{bmatrix} \quad (3.2)$$

In Figure 3.4 this operation is illustrated.

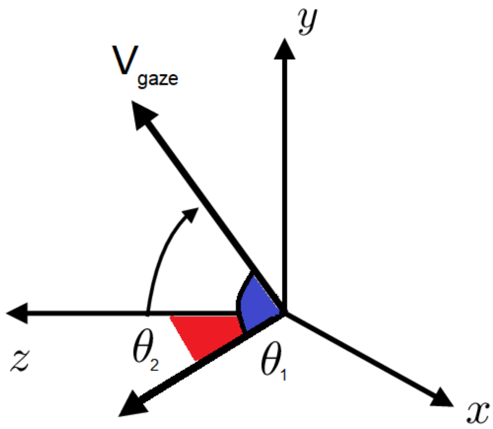


Figure 3.4: Rotating z axis to align with V_{gaze}

The Multiplication of these matrices would provide the gaze rotation matrix, R_{gaze}, as in equation 3.1.

$$R_{\theta_1} \cdot R_{\theta_2} = R_{gaze} \tag{3.3}$$

After computing the rotation matrix as R_{gaze}, inverse kinematic model is used to compute the values of the last three joints of the UR5, head group joints.

3.1.2 Head Gestures

In this chapter mathematical models for UR5 Collaborative robot's head behaviours will be presented. In general head behavior can be defined as several tilt action on the different angles that bisect the head from three orthogonal planes. These orthogonal planes are shown in Figure 3.6. Different head behaviors can be performed by moving the head on these planes.

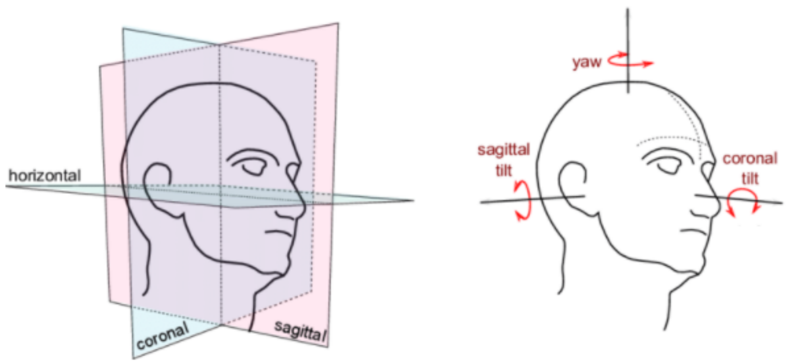


Figure 3.5: Human head Rotation axis

Consecutive top-down tilt on the sagittal plane would be our main approach for successful head nodding behavior. In other words, after a successful assembly task, this action can be used to inform human participants about the state of the job. Again, with the consecutive movement of the head on the horizontal plane an head shake behaviour will be implemented. Lastly, a head tilt behaviour is implemented by the rotating the head on coronal plane.

These head gestures can also be represented as rotational matrices for the robot's head Group. Like a animal or human head, the UR5 robot also has three different rotation planes. In Figure 3.6 these rotation axes are highlighted while the robot is focusing on an object, a grey box. Similar to gaze behavior, these head gestures can also be achieved with the inverse kinematic solutions of the rotational matrices for the head group joints. In Figure 3.6 from the initial position of the robot head group's state, rotations on three different axis would map to head shake, head nod, and head tilt behaviours.

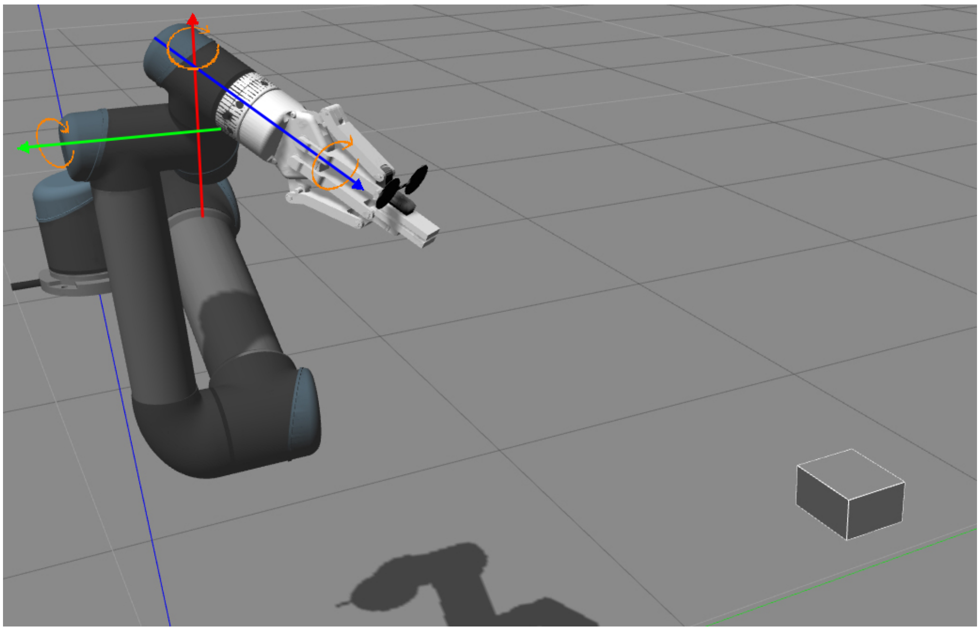


Figure 3.6: UR5 Robot Rotation axis, Blue z-axis, green y-axis and red is x-axis of the robot head group.

$R_{initial}$ represents the current rotation of the head group joints. In equation 3.4, $R_x(\alpha)$ represents the rotation movement in x-axis, in equation 3.5, $R_y(\beta)$ represents the the rotation movement in y-axis and lastly in equation 3.6 $R_z(\gamma)$ represents the the rotation movement in z-axis.

$$R_{initial} \cdot R_x(\alpha) \tag{3.4}$$

$$R_{initial} \cdot R_y(\beta) \tag{3.5}$$

$$R_{initial} \cdot R_z(\gamma) \tag{3.6}$$

More explicitly, these equation can be represented in 3.7, 3.7, 3.9.

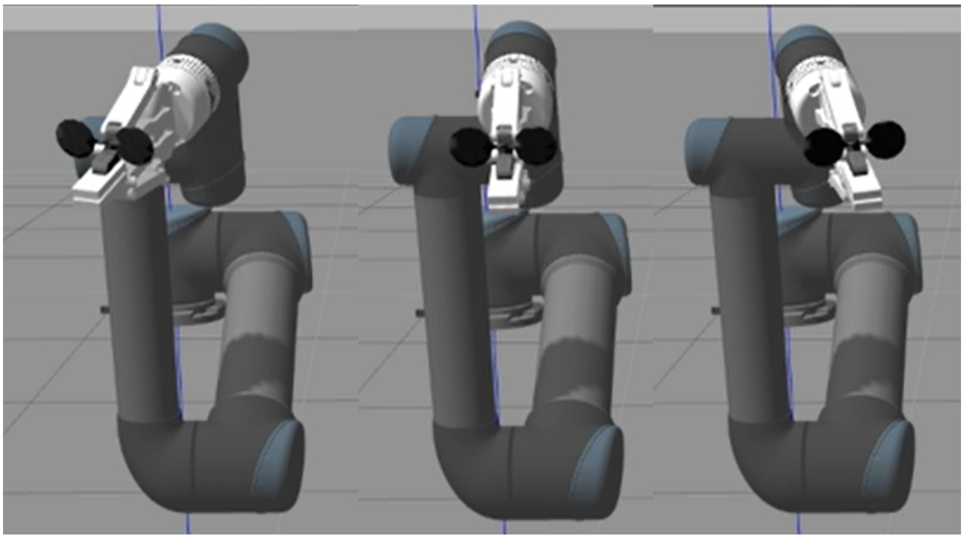
$$R_{initial} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{3.7}$$

$$R_{initial} \cdot \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \tag{3.8}$$

$$R_{initial} \cdot \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.9}$$

The inverse kinematic solution of equation 3.7 for the head group can achieve a horizontal plane movement for the end-effector. With similar logic different, movement of the head represented mathematically for our three head gestures.

3.2 Head Shake



(a) Left Head State. (b) Initial State. (c) Right Head State.

Figure 3.7: Head Shaking States.

In Figure 3.7 three different states of head shake illustrated. Respectively moving from initial to left head state, from left head state to right, and lastly going back to initial state would imitate a head shake behavior for UR5 Collaborative robot. These head states can also be represented as rotational matrices respectively, left head state in equation 3.10, initial head state in equation 3.11 and right head state in equation 3.12.

$$R_{initial} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \text{ where } \alpha \text{ is, } \pi/5 = \alpha \tag{3.10}$$

$$R_{initial} \tag{3.11}$$

$$R_{initial} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(-\alpha) & -\sin(-\alpha) \\ 0 & \sin(-\alpha) & \cos(-\alpha) \end{bmatrix}, \text{ where } \alpha \text{ is, } \pi/5 = \alpha \tag{3.12}$$

3.3 Head Nod

Like Head Shake behavior, head nodding behavior states can also be represented with rotational matrices for head group joints. We have implemented the head-nodding gesture by moving the gripper along the vertical axis using the data recorded from head nodding in humans (Liu, T., et al., 2012) and generated the gesture with similar velocity profiles as illustrated in 3.8.

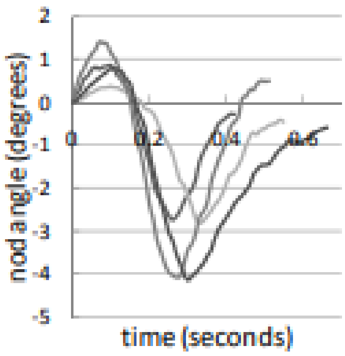
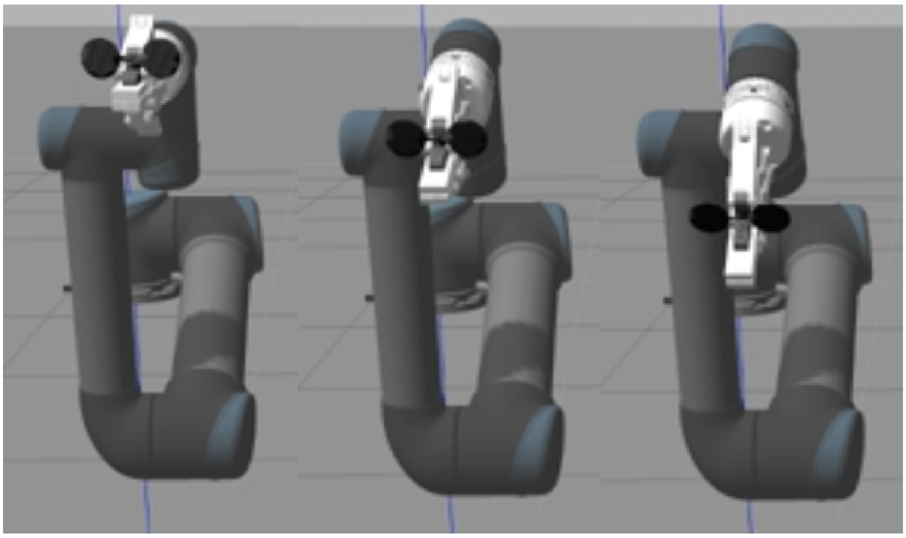


Figure 3.8: Human head movement graph for previously recorded head nodding behavior (Liu, T., et al., 2012).

For head nodding behaviour again there are three different states as illustrated in Figure 3.9. Consecutive movement, along with these states, would imitate a nodding behavior. In other words, respectively, going from Initial state to Upper Head State, from Upper Head State to Lower Head State, again going from Lower Head State to Upper Head state, and lastly going Upper Head State to Initial State would create a nodding behavior for UR5 Collaborative robot. These states similar to nodding behaviour can also represented in three different equations, respectively Upper Head State in equation 3.13, Initial Head State Equation in equation 3.14 and Lower Head State in equation 3.15.

$$R_{initial} \cdot \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \text{ where } \beta \text{ is, } \pi/180 = \alpha \tag{3.13}$$

$$R_{initial} \tag{3.14}$$



(a) Upper Head State. (b) Initial State. (c) Lower Head State.

Figure 3.9: Head Nodding States.

$$R_{initial} \cdot \begin{bmatrix} \cos(-\beta) & 0 & \sin(-\beta) \\ 0 & 1 & 0 \\ -\sin(-\beta) & 0 & \cos(-\beta) \end{bmatrix}, \text{ where } \beta \text{ is, } \pi/45 = \alpha \tag{3.15}$$

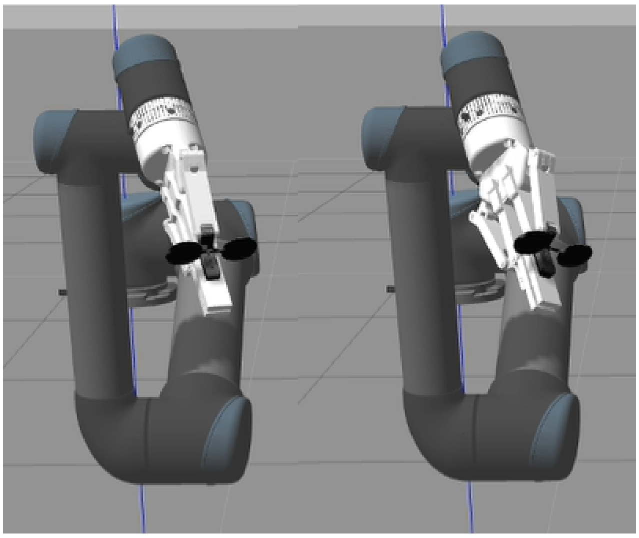
3.4 Head Tilt

Another mathematically modelled head gesture is head tilt. A small rotation on the third and final head plane, coronal plane, would create this behavior. Similar to other head gestures, mathematically this behavior can also be represented with rotation matrices as shown in equation 3.6. Tilt profiles used from previously tested humanoid robot experiments as 20 degree (Mara & Appel, 2015).

Again consecutive movement along these states would imitate a Head Tilt action. In this time, going from initial to coronal tilt state and going back to initial state would create this behavior. Initial state's mathematical representation shown in equation 3.16 and coronal tilt state's mathematical representation shown in equation 3.17.

$$R_{initial} \tag{3.16}$$

$$R_{initial} \cdot \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ where } \gamma \text{ is, } \gamma = \pi/12 \tag{3.17}$$



(a) Initial Head State. (b) Coronal Tilt.

Figure 3.10: Head Tilt Behaviour States.

3.4.1 Eye Contact Behaviour

Eye contact is another method that has been used behaviour experiments. This behavior is mainly used to create better, responsive interactions between robots and humans. It has been developed to respond to human eye contact. To achieve this V_{gaze} of the human is observed by the main code continuously and checked if it intersects with Robot Head Group. From the forward kinematics of the UR5 Collaborative robot, a position of 6^{th} joint is found, and a spherical area is generated to minimize possible instability coming from V_{gaze} vector of human. Since the human eye is difficult to track, unwanted mobility can occur while finding V_{gaze} thus; this spherical region creates an error gap. Mathematically this process can be represented as a line-sphere intersection equation in 3.18.

$$d = (b)^2 - 4 \cdot a \cdot c \tag{3.18}$$

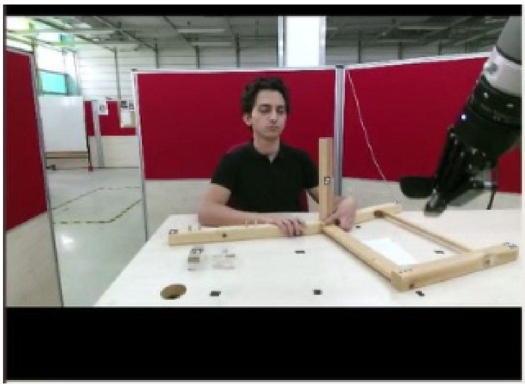
where a, b and c represented in equations respectively in Equation, 3.19, 3.20 and 3.21.

$$a = U \cdot U \tag{3.19}$$

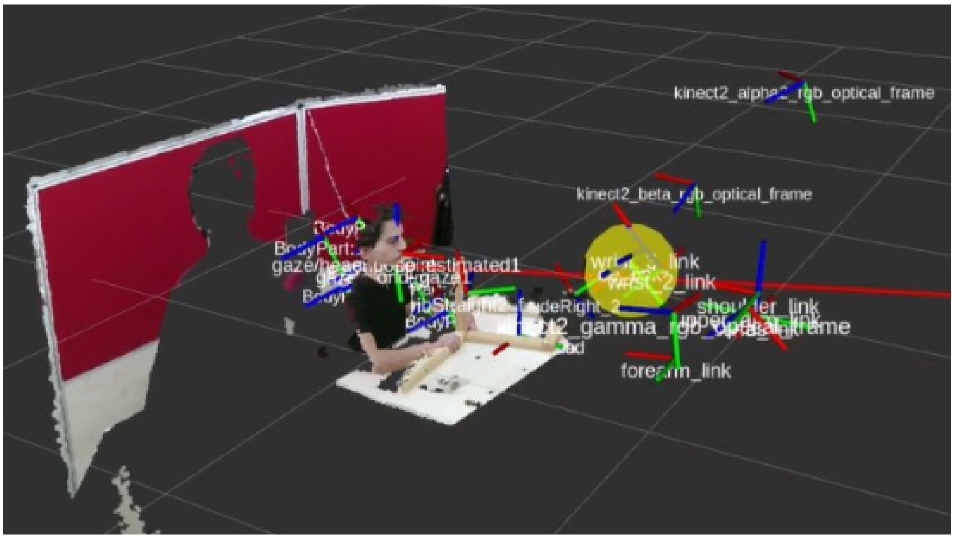
$$b = 2 \cdot U \cdot (P - C) \tag{3.20}$$

$$c = Q \cdot Q - (r \cdot r) \tag{3.21}$$

U represents the unit vector in the direction of V_{gaze}, P is a point in V_{gaze}, r is radius of the sphere, C is the centre of the sphere which is position of 6^{th} joint and lastly d in Equation 3.18 output. If d is higher or equal to 0, we can think that V_{gaze} is intersecting with the sphere. In other words, the human gaze is looking at the robot head. Moreover this sphere is visualized in Figure 3.11.



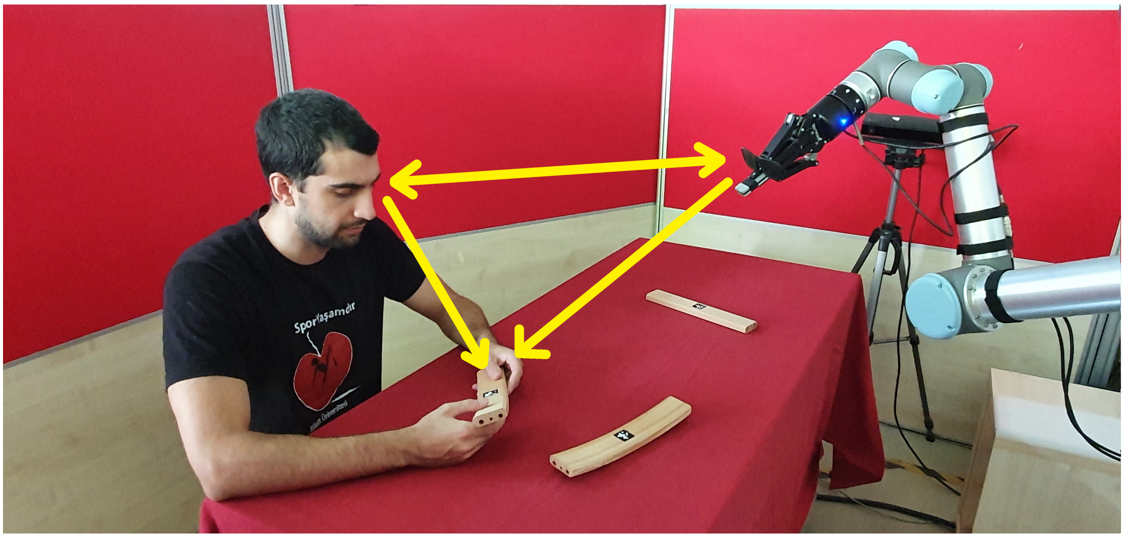
(a) Human Robot Environment.



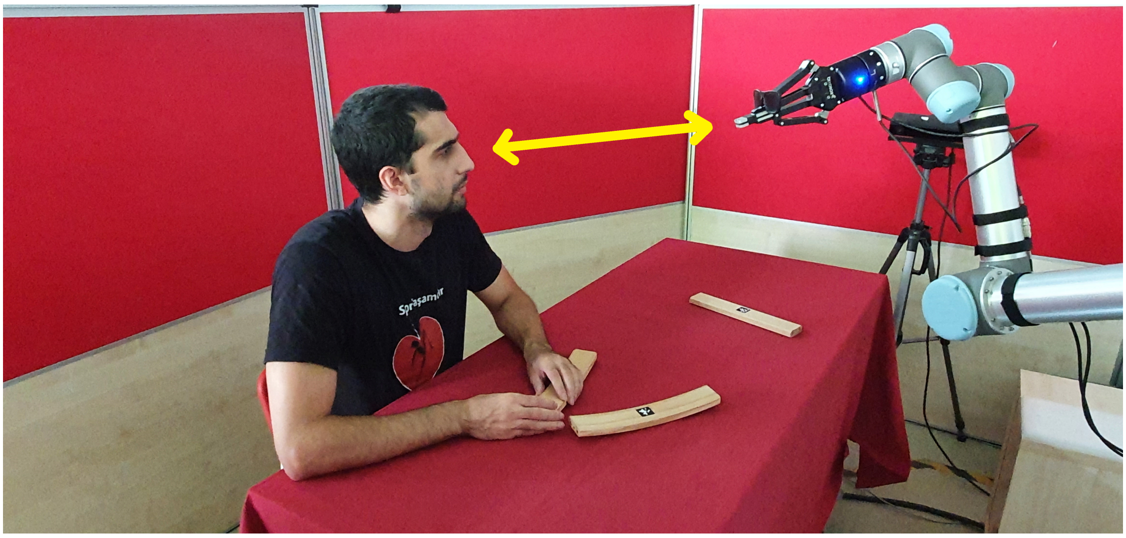
(b) Gaze Sphere visualization for Human-Robot Environment. When V_{gaze} of human participant, intersects with the yellow area eye contact behaviour is triggered.

Figure 3.11: Robot Gaze Sphere illustration

After human worker looks at the robot head, our gaze behaviour for robot is triggered with that UR5 Collaborative robot turns his head to human participant. This action is visualized in Figure 3.12.



(a) Looking at the Pieces.



(b) Looking at the Robot.

Figure 3.12: Shared and mutual gaze as implemented on UR5 cobot

3.5 Assembly Behavioural map

In this section we proposed a behavioral model for assembly task between UR5 Colaborative robot and human worker. These behaviors are developed while observing several other HRI experiments mentioned in the literature review. However, because of the UR5 Collaborative robot structure, not every single behavior can be implemented. For example, pointing an object with a finger is an effective way of shifting attention; however, UR5 Robot does not have a natural hand to point objects. Its end-effector could not mimic such behaviors. Moreover, there were many different robots with different capabilities in the literature, and to implement some of their behaviors on the UR5 Collaborative robot, intuitive models were used. UR5 Collaborative robot has six DOFs, and to perform these models on UR5, its structure is separated in half. The first three joints represented the body section, and the last three joints represented the head section. In other words, the UR5 collaborative robot body has modeled like a snake. Moreover, with the end-effector's natural look, a duck-like Figure is created. Similar design principles for the UR5 Collaborative robot were used previously and tested inside a simple assembly task (Terzioğlu et al., 2020)

Different behaviors were implemented on UR5 Collaborative robot for an assembly task. These behaviors can be listed below:

- **Gaze Behaviour** is implemented on the Head Group of UR5 Collaborative robot to track human actions and shift attention. Moreover, this behavior is used in different scenarios, such as mutual gaze.
- **Head Gestures** are implemented as different body language actions. Different behaviors, such as head nodding, head tilt, etc., are developed on the UR5 Collaborative robot, and with them, the feedback was given to workers about the assembly process.

With the proposed behaviors, a behavioral model developed for an assembly task. To build an end product in an assembly task, many different actions occur both from the human and robot sides. Additionally, a failure case can occur depending on the human worker's actions. In our behavioral map, these failures are addressed and, depending on the case, failures handled by a collaborative robot.

UR5 cobot actions are organized similar to a state machine, and depending on the inputs, robot behaviors were decided. This state machine can be seen in Figure 3.13.

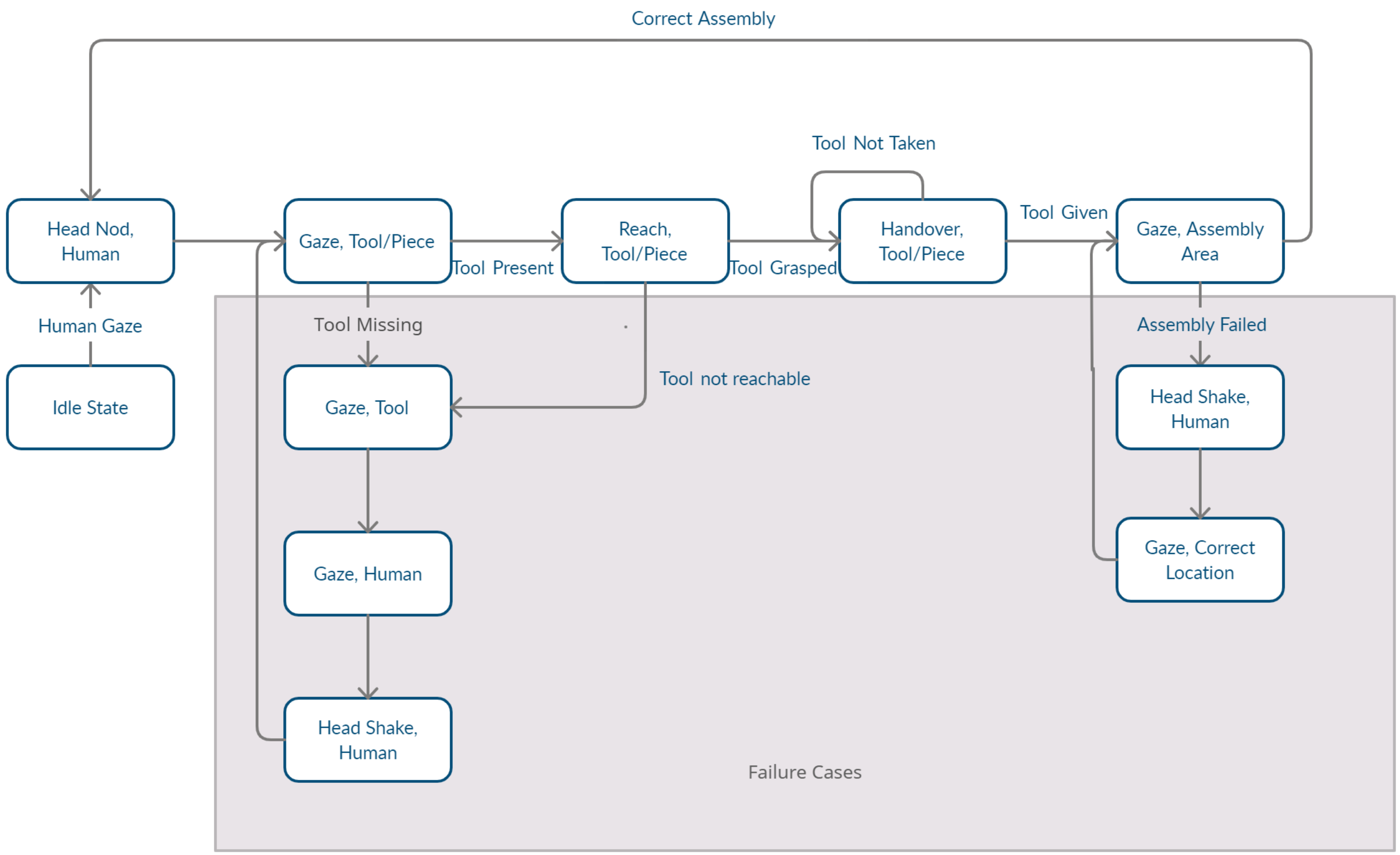


Figure 3.13: Assembly State Machine

Further explanation for actions illustrated in Figure 3.13 :

- **Idle** initial state where robot wait for input to be triggered.
- **Gaze action** is mapped to our gaze behavior proposed in this thesis. This action takes the variable written after the comma, as in Figure 3.13. This variable represents the robot head group direction while acting.
- **Head Shake action** is mapped to our head shake behavior that is proposed in this thesis. This action takes a variable after the comma. This variable also represents the robot head group direction.
- **Head Nod** is mapped to our head nodding behavior proposed in this thesis. This action takes a variable after the comma. This variable also represents the robot head group direction.
- **Reach action** is a recorded action that involves gripping the tool or the piece. After reaching action, the handover process follows.
- **Handover action** is also a recorded action. After reaching and gripping the assembly piece, it is given to the human worker.

Additionally, explanation for inputs:

- **Human Gaze** is a worker input that triggers the assembly process. It is detected after the mutual gaze.
- **Tool/Piece Missing** occurs when the required tool is missing in stock. It creates a failure case.
- **Tool/Piece Presents** input occurs when the tool/piece is available in the stock.
- **Tool/Piece Grasped** occurs when the required tool reached and grasped by the robot.
- **Tool/Piece not taken** happens when the worker does not take the tool/piece from the robot's manipulator.
- **Tool/Piece given** occurs when handover action finishes successfully.
- **Assembly Failed** occurred when the worker failed to combine assembly pieces.

UR5 cobot initially triggered by human worker gaze with the algorithm proposed in the mutual gaze section 3.4.1. Afterward, the UR5 cobot notifies the human worker by a head nodding gesture. At this point, the assembly process begins, and the robot gives a head cue to the tool or assembly piece to inform the human worker about the next action. If the necessary tool present in stock, the robot reaches the tool; otherwise, a failure case occurs. The UR5 robot tries to handle the error case, and firstly it directs its head to the worker and notifies the worker by head shake behavior. If the tool presents in the stock UR5, cobot tries to reach the tool by doing reach action. If the object is not reachable robot turns its head to the worker and notifies him by shaking its head.

If the robot could reach the tool/piece, it proceeds to the next state and grasps the piece/tool. Later, the robot starts the handover process and waits until the tool/piece is transferred to the worker. UR5 Cobot shifts the worker's attention when the tool is given by looking at the assembly area. At this stage, the worker starts the assembly by combining pieces. If everything goes in order, the loop goes to the beginning and checks whether the assembly process is finished.

CHAPTER 4

EXPERIMENT AND RESULT

4.1 Experimental Framework

In this section experiment setup which consist of consist of an UR5 Robot, a Layer 2 switch, a Kinect camera and PC with Ubuntu 16.04 (Kinetic Ros installed) operating system. The main architecture of the setup illustrated in the below Figure 4.1:

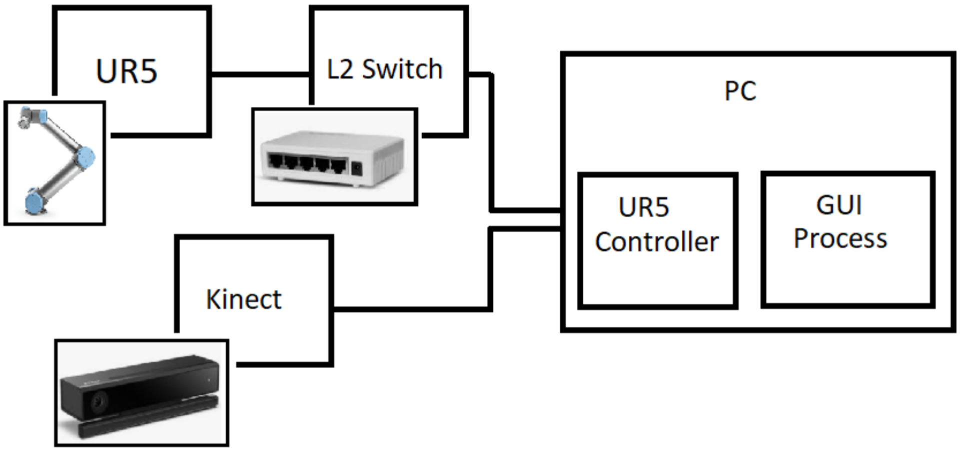


Figure 4.1: Experiment Architecture visualized

4.1.1 Robotic Platform

6-DOF, UR5 Collaborative robot was used in experiments for this thesis. It was elevated to the table level through a static platform in order to support a person who sit in front of a table. Additionally robot control panel used to stop the robot in an emergency case. Since the robot planner outputs are not monitored, unexpected or unwanted behaviours might occurred. Experiment environment illustrated in Figure4.2.

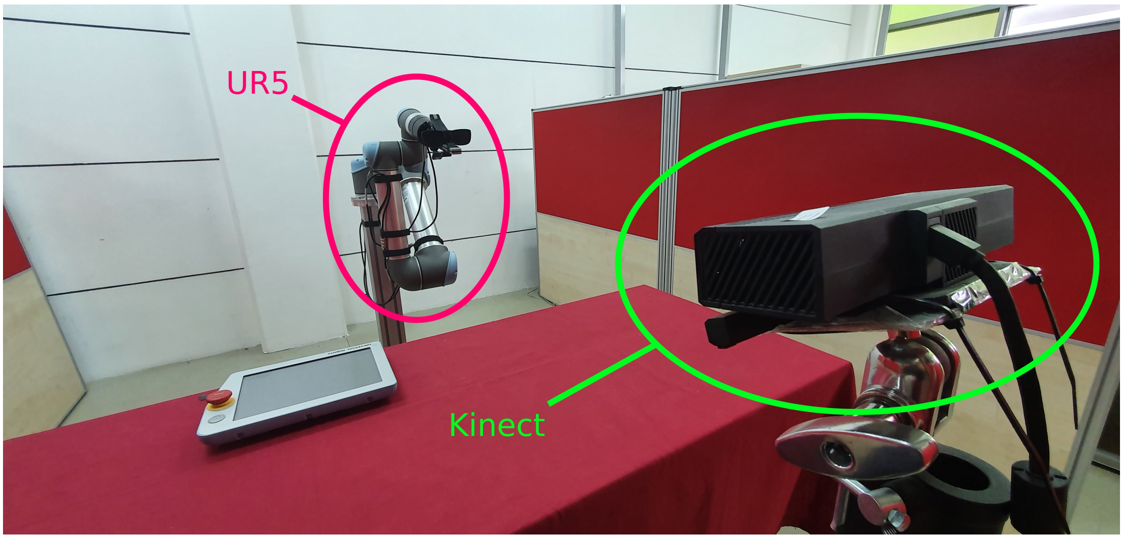


Figure 4.2: General overview of the experiment environment.

4.1.2 Kinect Camera

The Kinect camera, Figure4.3, is a multi functional camera developed for a commercial video game console: however it is also widely used among different robotic and perception experiments. It features a depth sensor and an RGB camera with different resolution options. Furthermore, in our experiment Kinect represents the human perception and it stands on a tripod to track robot actions.

Figure 4.3: Kinect camera visual.

4.1.3 Controller Computer

A computer was connected to both Kinect camera and the robot. UR5 Collaborative robot was connected to this computer through a switch. With this computer, robot actions planned and executed with the inputs coming from user interface. Moreover to control robot actions ROS framework used. ROS framework is a set of libraries and programs to develop robotic applications.

Inside this computer Zoom application is running for participants to remote control. Zoom is a free teleconference and remote control tool which is widely used. With this application participants could be able to remotely access to the experiment environment.

4.1.4 Python ROS and User Interface

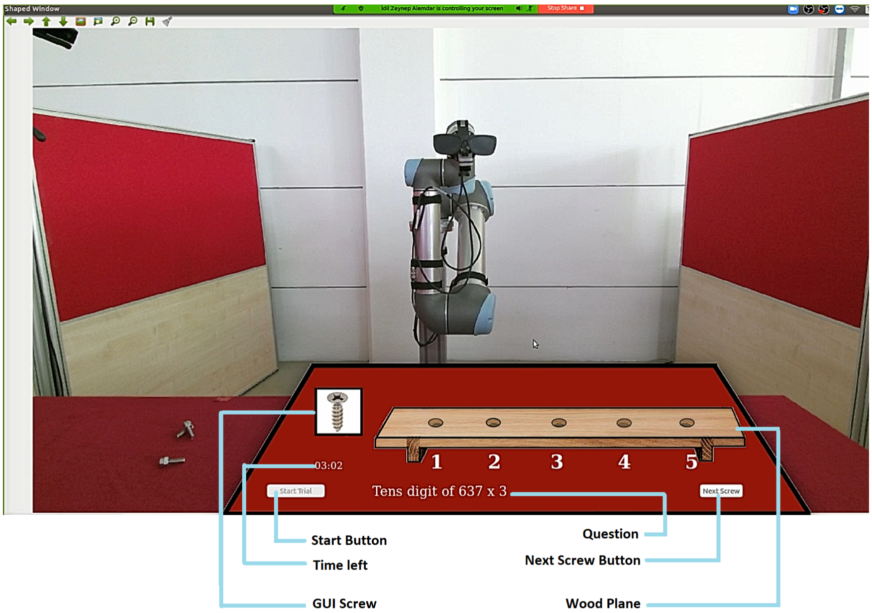


Figure 4.4: Experiment GUI Explained

In order to generate and design a virtual assembly table, Wxpython GUI library (Team, 2019) is used. GUI and the main code which controls the robot(again through python language), run in two different processes. UR5 Robot plan execution is a blocking process and because of the GIL in python, both of these codes could not run in different threads. Furthermore ROS framework does not allow different forked processes to control the robot. Because of these reasons and since there were no strict performance limitations, both processes communicated through a UDP socket interface which is relatively slower than some other IPC methods. With these techniques UR5 robot's planing and human actions in GUI, could be processed in parallel. Additionally UR5 Collaborative robots have an internal cue which put an order to the planned actions. In other words, you can send several different actions in a short duration of time, however UR5 will do all of these actions by order with out interfered by another action.

When the participants interacted with the robot through the GUI, socket interface was triggered and sending string commands i.e "HeadShake" which commands robot to shake its head. There were several different string commands implemented:

- **"HeadShake"** commands robot to initiate head shaking behaviour.
- **"HeadNodding"** commands robot to initiate head nodding behaviour.
- **"GazeAversion"** robot joint distance gets smaller and robot looks to the ground.
- **"BringScrew"** commands robot to bring the screws.

- **"Gaze"** commands robot to gaze at a certain point which is also send after the string command, in 3D space .

Along with these behaviours, other string commands are also implemented for head tilt, head nodding etc. However these behaviours were not used in this experiment.

4.2 Experiment

4.2.1 Experiment Methodology

The experiment had a within-subject design with three conditions. Each condition represented the degree of correctness in robot's gaze-pointing gesture. More specifically, the robot provided a gaze cue to the participants. The task of the participant was to drive the screw into the correct hole. The information about the correct hole location was given as a challenge to the participant (a multiplication of two two-digit numbers). In each trial, the participant calculated the multiplication to find out the correct location of the hole and drive the screw to the hole. The robot pointed the hole in different degrees of correctness in the three experimental conditions, as presented in more detail below.

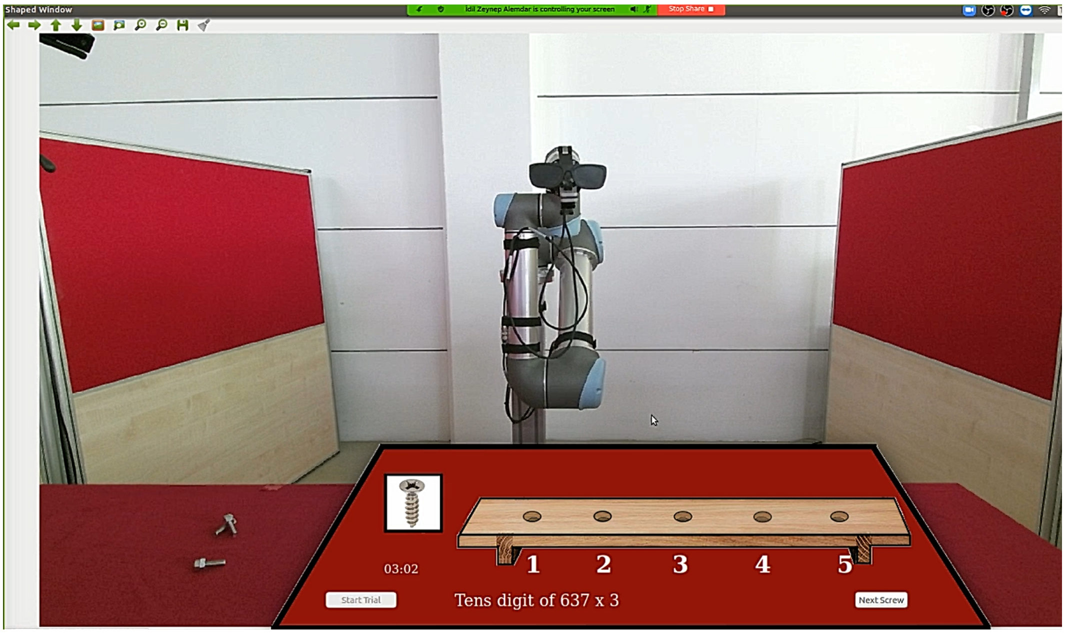
The participants were given 200 seconds in each condition. The order of the conditions was randomly distributed across the participants. The task of the participant was to drive as many screws as possible into the correct hole. The presence of the multiplication task made it a challenging task to find the correct hole to drive the screw due to cognitive load (Sweller, 1988). The difficulty of the multiplication was similar across the questions. The participants calculated the tens digit of multiplication to identify the location of the correct hole. In each question the second multiplier remained the same where the initial three digits were changed to map all five holes on the wood pane. In Figure 4.5a there is an example multiplication question as "Tens digit of 637×3".

The gaze-pointing of the robot (to the hole location) provided an opportunity to the participants for interaction. In each trial the robot gaze at one of the holes to give a gaze cue to the participant. Nevertheless, the robot pointed the holes with varying correctness as a function of the experimental condition. Thus, we designed the experiment's independent variable by manipulating the "trust factor" with three conditions (%0, %50,and %100). If the trust factor were %0, the robot would indeed gaze at an incorrect hole. Similarly, it would always gaze at the correct hole when the trust factor was %100. Finally, it was a fifty-fifty chance to gaze at the correct hole to the participant when the trust factor was %50. The participants were not informed about the value of the trust factor in the instructions session. Instead, they were informed that the robot would time-to-time point at a correct hole.

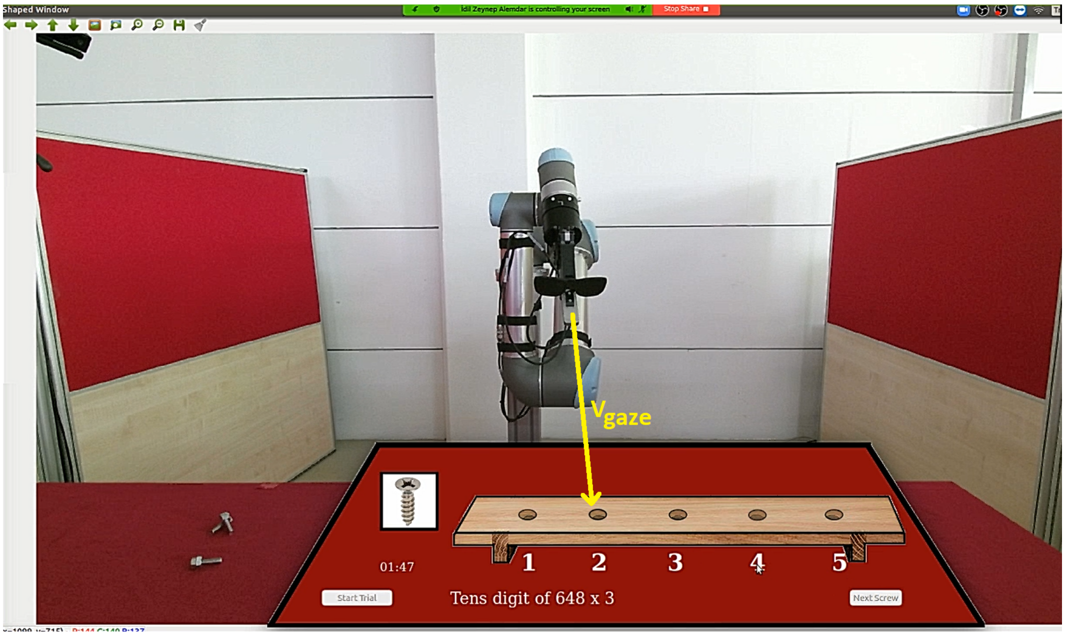
In order to progress in the trials, the participant clicked with their mouse on the holes to drive a screw. After the screw placed, they were able to proceed with the next screw by pressing the "Next Screw" button at the right corner of the GUI (see 4.2). The "Next Screw" button became clickable when the participant drove the screw into the hole. When the “Next Screw" button was pressed, the interface played a positive sound if the screw was driven into the correct hole. The participant heard a negative sound in case of having driven the screw into an incorrect hole. The robot actions are presented in Table 4.1.

Table 4.1: Experiment Action table

Cobot guidance	Human action	Cobot Response
Gaze at correct hole	Driving screw to Correct Hole	No action
Gaze at correct hole	Driving screw to Wrong Hole	Gaze down and head shake
Gaze at wrong hole	Driving screw to Correct Hole	No action
Gaze at wrong hole	Driving screw to Wrong Hole	Gaze down and head shake



(a) Robot looking directly to camera



(b) Robot looking to virtual hole number 2

Figure 4.5: Experiment Environment

4.2.2 Questionnaires

After the administration of each experimental condition, the participants were asked to evaluate the session by filling in two questionnaires. The first one is the "Godspeed Questionnaire" by Bartneck (Bartneck, Kulić, Croft, & Zoghbi, 2009), which evaluates Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of robotic systems using five items for each measure. Participants used a ten-point rating scale to express agreement or disagreement with the items. For the second questionnaire, eight measures were chosen from the "Toolkit for Measuring Acceptance of an Assistive Social Robot" by Heerink (Heerink et al., 2009): Anxiety, Attitude, Perceived Adaptability, Perceived Enjoyment, Perceived Sociability, Perceived Usefulness, Social Presence, and Trust. Lastly, the participants were asked to comment about the robot's character for each corresponding trial. The participants also filled in a TIPI personality questionnaire at the end of the experiment (Gosling, Rentfrow, & Swann Jr, 2003).

4.3 Data analysis and Results

This section presents the method of analysis and the preliminary results of the experiment. We collected data from 19 participants (six female, mean age M = 24.84, SD = 4.87). In general all participants were between age of 20 and 27 except two participants (Their ages were 41 and 30).

The performance of the participants was measured in terms of the number of driven screws per experimental condition, the average duration from the onset of the screw display on the screen to driving it into the hole, rate of correctly driven screws and the average duration before clicking the "Next Screw" button. In our case independent variable was trust factor which changed in every experiment trial as %0 (T0), %50 (T50) and %100 (T100) percent and its effect on participants were observed. To analyze the questionnaire results and the performance values we used one-way ANOVA for three different trust factor.

Initial results indicates an improvement in the task performance in terms of the number of correctly driven screws, when the trust factor goes from %0 (M = 20.10, SD = 7.60) to %50 (M = 20.78, SD = 7.09) and %100 (M = 24.73, SD = 7.64). We have also obtained a similar pattern in the duration of task accomplishment. The participants spent the longest time (M = 5.35, SD = 2.70) under the lowest trust factor, a shorter time under the %50 trust factor (M = 6.31, SD = 2.41), and the shortest time (M = 6.66, SD = 2.71) with the highest trust factor. Interestingly a descriptive analysis of the answers to the questionnaires shows that some of the participants reported that the robot tried to trick them when the trust rate was %50. To sum up, participants' performance improved with increasing trust ratio. The highest accuracy values were achieved when the robot gaze-pointed the correct hole with %100 rate of correctness (i.e., the fully trustable robot). There were statistically significant differences between group means for correct screw rates as determined by one-way ANOVA (F(2,36) = 5.368, p = .009, η^2 = .230) and total driven screw numbers (F(2,36) = 4.067, p = .026, η^2 = .184) as illustrated in Figure 4.6. Post-hoc analysis also showed significant difference between T100, with T50 (p=0.061) and T0(p=0.036) cases for correct screw rate. Again for total screw number, there was a significant difference between T100 with T0(p=0.009). Average screw times (F(2,36) = 2.355, p = .109 η^2 = .116) and average time for "Next Screw" button click (F(2,36) = 1.567, p = .223 η^2 = .080) analysis didn't show any significance.

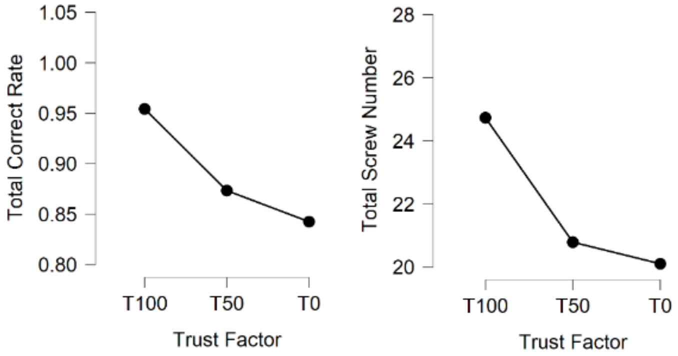


Figure 4.6: Total attached screw number and correctly attached screw ratio for three different trust factor.

Participants also tend to approach skeptical at the beginning of each condition. In Figure4.7, it was shown that the trust bonding process of participant no 17. Screw drive times start in between 7-8 seconds and drops with the correctly driven screws.

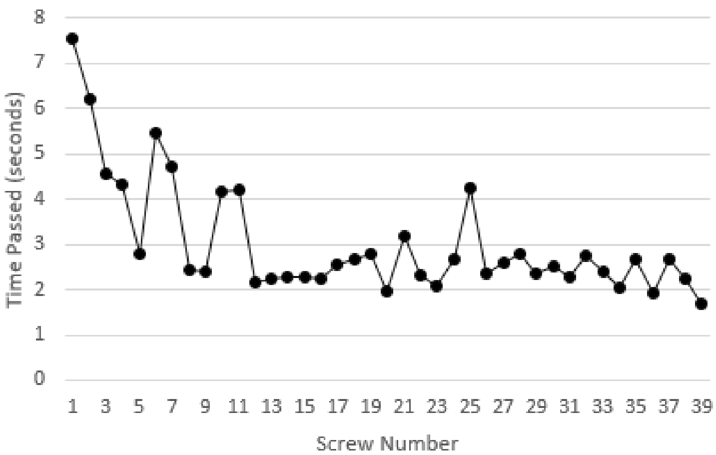


Figure 4.7: Participant 17 trust bonding process for Trust factor %100

These results are also supported by the questionnaires which indicates that the robots actions were trusted by the participants significantly (F(2,36) = 53.011, p < .001) as shown in Figure 4.8. Post-hoc analysis was between T100 with T0 (p<.001) and T50 (p<.001) demonstrated that the human choose not to trust the unreliable robot.

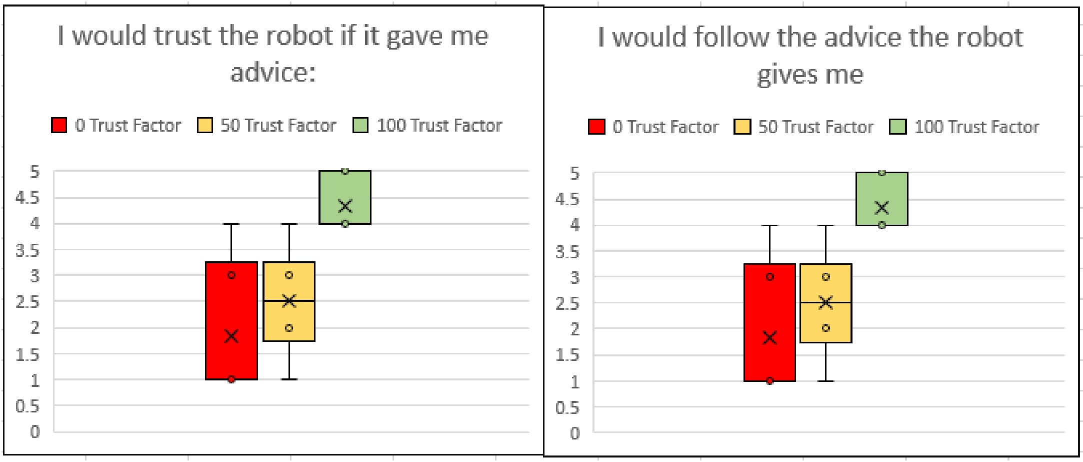


Figure 4.8: Questionnaire results for Trust attribution.

These findings may indicate that the participants interacted more intensively when the robot gaze-pointed correct locations, even partially. One may interpret the findings as humans' tendency for robot interaction under sub-optimal conditions (i.e., a partially trustable robot). The questionnaire results suggest that a partially trustable robot may provide a better sense of human-likeness than a fully deterministic robot. There was a significant difference in machine-like and human-like analysis ($F(2,36) = 3.543$ $p < .039$) between T0 and T50 ($p < 0.45$). Additionally fully trustable robot with %100 trust factor, rated as sensible ($F(2,36) = 4.312$ $p < .021$ $\eta^2 = .193$), competent ($F(2,36) = 17.971$ $p < .001$ $\eta^2 = .500$) , friendly ($F(2,36) = 11.320$ $p < .001$ $\eta^2 = .386$), knowledgeable ($F(2,36) = 17.742$ $p < .001$ $\eta^2 = .496$), intelligent ($F(2,36) = 24.016$ $p < .001$ $\eta^2 = .572$), responsible ($F(2,36) = 18.467$ $p < .001$ $\eta^2 = .506$), pleasant ($F(2,36) = 7.876$ $p < .001$ $\eta^2 = .304$), nice ($F(2,36) = 9.282$ $p < .001$ $\eta^2 = .340$) and likeable ($F(2,36) = 9.58$ $p < .001$ $\eta^2 = .359$). Additionally Post-hoc results of these questionnaires for T100 case also gave significant difference ($p<0.05$) over others.

We included the each question results in Figure 4.9.

Figure 4.9: "Godspeed questionnaire" results represented with box analysis.

Questionnaire results are also analyzed in constructs of Trust, Anxiety, Attitude, Perceived Adaptability, Perceived Enjoyment, Perceived Sociability, Perceived Usefulness, Social Presence (Heerink et al., 2009) as in Figure 4.11 and Anthropomorphism, Animacy, Likeability, Perceived Intelligence, Perceived Safety (Bartneck, Kulic, & Croft, 2009) as in Figure 4.10 . In our findings, several constructs had significance, such as: Anthropomorphism ($F(2,36) = 11.698$, $p < .001$, $\eta^2 = .394$), Animacy ($F(2,36) = 11.698$, $p = .031$, $\eta^2 = .176$), Likeability ($F(2,36) = 11.698$, $p = .001$, $\eta^2 = .441$) in "Godspeed Questionnaire" (Bartneck, Kulic, & Croft, 2009). Additionally for the questionnaire "Toolkit for Measuring Acceptance of an Assistive Social Robot", constructs such as: Trust ($F(2,36) = 64.294$, $p < .001$, $\eta^2 = .781$), Anxiety ($F(2,36) = 15.589$, $p < .001$, $\eta^2 = .464$), Attitude ($F(2,36) = 15.812$, $p < .001$, $\eta^2 = .468$), Perceived Adaptability ($F(2,36) = 18.087$, $p < .001$, $\eta^2 = .501$), Perceived Enjoyment ($F(2,36) = 14.023$, $p < .001$, $\eta^2 = .438$), Perceived Sociability ($F(2,36) = 17.131$, $p < .001$, $\eta^2 = .488$) and Perceived Usefulness ($F(2,36) = 19.993$, $p < .001$, $\eta^2 = .526$), had also significance.

"Godspeed Questionnaire" post-hoc analysis for Anthropomorphism T50 and T100 cases significantly differed from the T0 case (with $p < .001$). For Animacy, T100 had a significant difference compared T0 case (with $p = .030$), and lastly, for Likeability, T100 and T50 cases had significant differences compared to T0.

In the "Measuring acceptance of an assistive social robot: a suggested toolkit" questionnaire results, for Trust, the T100 case had a significant difference compared to T50 and T0 cases (with $p < .001$ for both). For Anxiety T0 case had a significant difference compared to the T100 case (with $p < .001$). For Attitude T100 case had a significant difference compared to the T0 case, for Perceived Adaptability T100 case had a significant difference compared to T50 and T0 cases. For Perceived Enjoyment, Perceived Usefulness, and Perceived Sociability, each case has a significant difference compared to every other case.

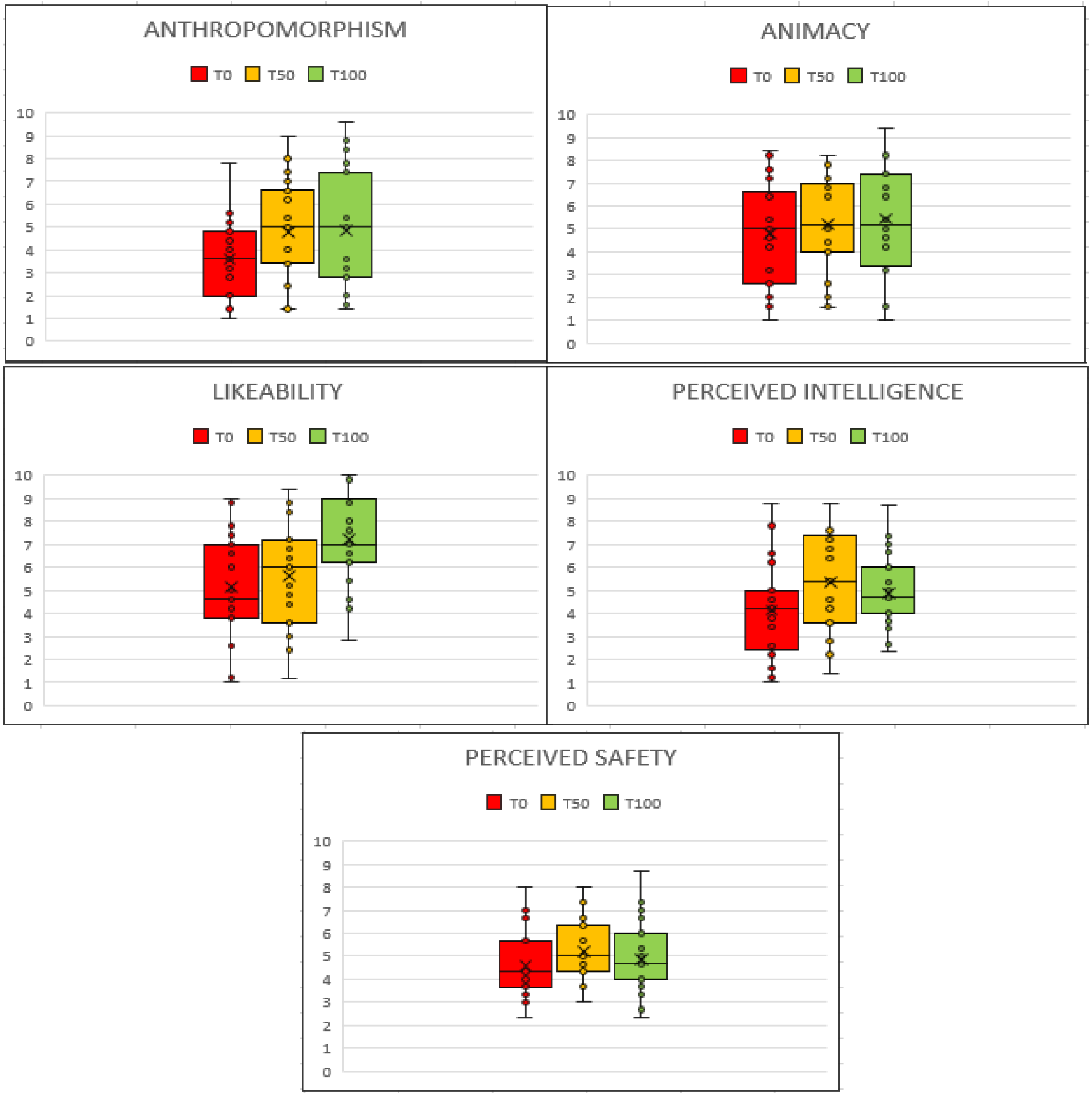


Figure 4.10: Questionnaire results for "Measurement instruments for the anthropomorphism, animacy likeability, perceived intelligence, and perceived safety of robots" (Bartneck, Kulic, & Croft, 2009).

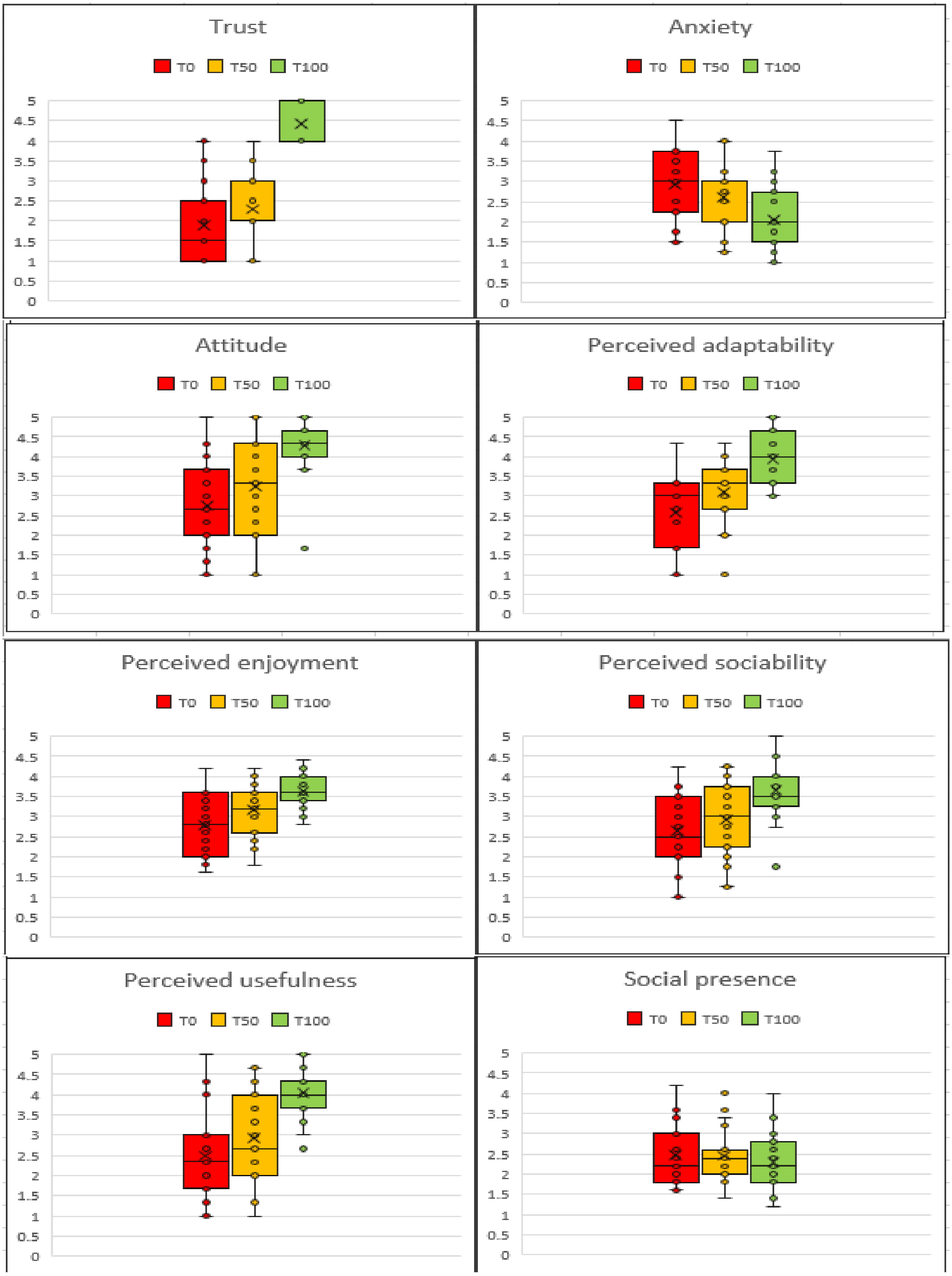


Figure 4.11: Questionnaire results for "Measuring acceptance of an assistive social robot: a suggested toolkit" (Heerink et al., 2009)

CHAPTER 5

DISCUSSION AND CONCLUSION

Compared to other traditional robotic systems, which come with heavy safety equipment and rules, we perceive cobots not as inanimate manipulators that carry out manipulative tasks alongside humans but as interactive workmates for the human workers. They can support people to refine the collaborative task and make it easier for them. They play a role in the quality processes, maintenance, and assembly lines of a product parallel with humans. In our case, this role was remembering the assembly process guiding human workers.

We argued, the use of non verbal behaviors, such as gazing, head nodding or shaking can be deployed to improve the quality of HRI with minimum hindrance on the robot's functionality and reliability in industrial settings. These gestures were adapted from different robots through out the literature and they were automated through inverse kinematic models and effects of gazing and head shaking tested on a telepresent robotic experiment setup.

Telepresent robotic is a newly emerged field that allows operators to participate in remote locations through live video feeds. Due to the COVID-19 pandemic, we designed a telepresent experimental setup to conduct HRI experiments with subjects remotely and investigated trust attribution in cobots that use non-verbal behaviors for HRI. The setup consisted of the cobot platform placed in front of a table, faced by a Microsoft Kinect camera(used only as a RGB camera). The participant interacted with the cobot remotely via the Zoom teleconferencing system. The participant's task is to drive the maximum number of screws into the holes on a virtual wood plank within 200 seconds. The correct spot was given as a challenge via the multiplication of two numbers. In each trial, the participant calculated the multiplication to find out the correct location of the hole and drive the screw to the hole. The cobot provided a gaze cue to the participants to find out the trust attribution of the UR5 Collaborative robot. This methodology is considered as cognitive load and it is used to amplify the findings on trustworthiness.

Trust has been studied in various disciplines (including HRI and cognitive science) to comprehend the relationships between humans or between human and agents. The diversity lead to various definitions and theories for trust (Adams, Webb, & Bryant, 2001). Based on the results obtained from the conducted experiments, we conclude that there is a systematic relation between robot trust and participants' performance. In our experiments, we observe that when the robot reliability improved, the participants' task scores also improved slightly, going from %0 percent trust rate to %50 percent. However, when the robot became fully reliable, participants' performance and robot trustworthiness improved significantly. Additionally, people tend to feel like their interactions are more human-like when robot made mistakes sometimes. In %50 trust factor case robot received the highest human-likeliness scores compared to other trust factors. From the open-ended questions that were designed to understand what people think in general for the robot behaviour, we concluded that there were several different groups who interpreted the semi-trustable robot. Some people thought that the robot was doing mistakes intentionally to confuse the participants and some thought robot looked more human-like when it made mistakes. However, the robot with %100 trust factor received the highest scores for fields such as: Trust, Attitude, Perceived Adaptability, Perceived Sociability, Perceived Enjoyment Perceived Usefulness and Likeability.

In order to understand the implications of these results, future studies could increase the number of participants. Due to Covid-19 we had to change the experiment structure several times to adapt current conditions. This reduced the time for finding participants and increased the time for developing the experiment environment. Additionally because of the individual difference of the participants, experiment findings had high variance. In the future we are considering procedures like digit span task, to quantify the individual characteristics of participants. Also for the trust formation, time series analysis is planned which will improve the understanding of the results.

Lastly there is another mathematical method for gaze behavior which needs to be addressed. Instead of finding a rotational matrix for the last three joints of the UR5 robot, a virtual prismatic joint could be attached to the end of the last joint of the UR5 Robot inverse kinematic chain. With that configuration, there won't be and dependency for a gaze vector. Instead, one can find the inverse kinematic solution of a point location, P_{target}, and orientation, R_{target}, as in Equation 5.1, for the last four joints (Head group joints and newly added virtual prismatic joint). The gaze orientation would be equal to the orientation of the world frame. In other words, the reference frame and the point would be the gaze location in 3D space.

$$R_{target} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_{target} = \begin{bmatrix} P_{target_x} \\ P_{target_y} \\ P_{target_z} \end{bmatrix} \tag{5.1}$$

REFERENCES

Adams, B. D., Webb, R. D., & Bryant, D. J. (2001). *Trust in teams: literature review*. Humansystems Incorporated.

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *Journal of Human-Robot Interaction*, *6*(1), 25. doi: 10.5898/jhri.6.1.admoni

Andrist, S., Mutlu, B., & Gleicher, M. (2013). Conversational gaze aversion for virtual agents. *Intelligent Virtual Agents Lecture Notes in Computer Science*, 249–262. doi: 10.1007/978-3-642-40415-3_22

Andrist, S., Mutlu, B., & Tapus, A. (2015). Look like me. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. doi: 10.1145/2702123.2702592

Andrist, S., Tan, X. Z., Gleicher, M., & Mutlu, B. (2014). Conversational gaze aversion for humanlike robots. *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. doi: 10.1145/2559636.2559666

Bartneck, C., Kulic, D., & Croft, E. (2009). Measurement instruments for the anthropomorphism, animacy likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*(1), 71–81.

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

Botsman, R. (2017). *Who can you trust?: How technology brought us together and why it might drive us apart*. Hachette UK.

Breazeal, C., Kidd, C., Thomaz, A., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. doi: 10.1109/iros.2005.1545011

Breazeal, C., & Scassellati, B. (n.d.). How to build robots that make friends and influence people. *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients (Cat. No.99CH36289)*. doi: 10.1109/iros.1999.812787

CarnesMS, M. B. (2019, Oct). *Head tilt in dogs: A clinical approach.* Retrieved from https://todaysveterinarypractice.com/head-tilt-in-dogs-a-clinical-approach/#:~:text=Theprimaryreasonfora,andlimbsrelativetogravity.

Cassell, J. (2000 April). *Embodied conversational interface agents* (No. 34). Communications of the ACM.

Cha, E., Kim, Y., Fong, T., & Mataric, M. J. (2018). A survey of nonverbal signaling methods for non-humanoid robots.
doi: 10.1561/9781680834093

Colgate, E., Wannasuphoprasit, W., & Peshkin, M. (1996). Cobots: robots for collaboration with human operators. In *Proceedings of the 1996 asme international mechanical engineering congress and exposition* (pp. 433–439).

Collaborative robotic automation: Cobots from universal robots. (n.d.). Retrieved from `https://www.universal-robots.com/`

Consistency in early vocabulary. (2021). *Variability and Consistency in Early Language Learning*. doi: 10.7551/mitpress/11577.003.0009

Cook, M. (1977). Gaze and mutual gaze in social encounters: How long—and when—we look others" in the eye" is one of the main signals in nonverbal communication. *American Scientist*, *65*(3), 328–333.

Costa, M., & Bitti, P. E. R. (2000). Face-ism effect and head canting in ones own and others photographs. *European Psychologist*, *5*(4), 293–301. doi: 10.1027//1016-9040.5.4.293

Darwin, C. (n.d.). Means of expression in animals. *The Expression of the Emotions in Man and Animals*, 83–115. doi: 10.1017/cbo9781139833813.005

Delaunay, F., & Belpaeme, T. (2012). Refined human-robot interaction through retro-projected robotic heads. *2012 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. doi: 10.1109/arso.2012.6213409

Emery, N. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience Biobehavioral Reviews*, *24*(6), 581–604. doi: 10.1016/s0149-7634(00)00025-7

Exline, R., Gray, D., & Schuette, D. (1965). Visual behavior in a dyad as affected by interview content and sex of respondent. *Journal of Personality and Social Psychology*, *1*(3), 201–209. doi: 10.1037/h0021865

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*(1). doi: 10.1186/s12984-020-0642-5

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the big five personality domains. *Journal of Research in Personality*, *37*, 504-528.

Halberstadt, A. G., & Saitta, M. B. (1987). Gender, nonverbal behavior, and perceived dominance: A test of the theory. *Journal of Personality and Social Psychology*, *53*(2), 257–272. doi: 10.1037/0022-3514.53.2.257

Heads, W. D. T. T., Llera, R., & Llera, C. R. (n.d.). *Why dogs tilt their heads.* Retrieved from `https://vcahospitals.com/know-your-pet/why-dogs-tilt-their-heads`

Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2009). Measuring acceptance of an assistive social robot: a suggested toolkit. In *Ro-man 2009-the 18th ieee international symposium on robot and human interactive communication* (pp. 528–533).

Hoque, M. M., Onuki, T., Kobayashi, Y., & Kuno, Y. (2013). Effect of robot's gaze behaviors for attracting and controlling human attention. *Advanced Robotics*, *27*(11), 813–829. doi: 10.1080/01691864.2013.791654

Iverson, J. M., Tencer, H. L., Lany, J., & Goldin-Meadow, S. (2000). The relation between gesture and speech in congenitally blind and sighted language-learners. *Journal of Nonverbal Behavior*, *24*(2), 105–130. doi: 10.1023/a:1006605912965

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1), 78–100. doi: 10.1037/0033-2909.100.1.78

Langer, A., Feingold-Polak, R., Mueller, O., Kellmeyer, P., & Levy-Tzedek, S. (2019). Trust in socially assistive robots: Considerations for use in rehabilitation. *Neuroscience & Biobehavioral Reviews*, *104*, 231–239.

Lasseter, J. (1987). Principles of traditional animation applied to 3d computer animation. In *Acm siggraph computer graphics* (Vol. 21, pp. 35–44).

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, *77*, 23–37. doi: 10.1016/j.ijhcs.2015.01.001

Li, Z., & Mao, X. (2012). Emotional eye movement generation based on geneva emotion wheel for virtual agents. *Journal of Visual Languages Computing*, *23*(5), 299–310. doi: 10.1016/j.jvlc.2012.06.001

Liu, C., Ishi, C. T., Ishiguro, H., & Hagita, N. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI 12*. doi: 10.1145/2157689.2157797

Liu, C., T., I. C., Hiroshi, I., & Norihiro, H. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. international journal of humanoid robotics. In *2012 7th acm/ieee international conference on human-robot interaction (hri)* (pp. 285–292).

Mara, M., & Appel, M. (2015). Effects of lateral head tilt on user perceptions of humanoid and android robots. *Computers in Human Behavior*, *44*, 326–334. doi: 10.1016/j.chb.2014.09.025

Mehrabian, A. (1968). Relationship of attitude to seated posture, orientation, and distance. *Journal of Personality and Social Psychology*, *10*(1), 26–30. doi: 10.1037/h0026384

Moore, & Dunham. (2014). Joint attention: Its origins and role in development. *Psychology Press*.

Morency, L.-P., Sidner, C., Lee, C., & Darrell, T. (2007). Head gestures for perceptual interfaces: The role of context in improving recognition. *Artificial Intelligence*, *171*(8-9), 568–585. doi: 10.1016/j.artint.2007.04.003

Mori, M. (1993). The uncanny valley. , *7*(4), 33-35.

Mwangi, E., Barakova, E. I., Díaz-Boladeras, M., Mallofré, A. C., & Rauterberg, M. (2018). Directing attention through gaze hints improves task solving in human–humanoid interaction. *International Journal of Social Robotics*, *10*(3), 343–355. doi: 10.1007/s12369-018-0473-8

Nao the humanoid and programmable robot: Softbank robotics. (n.d.). Retrieved from `https://www.softbankrobotics.com/emea/en/nao`

Normoyle, A., Badler, J. B., Fan, T., Badler, N. I., Cassol, V. J., & Musse, S. R. (2013). Evaluating perceived trust from procedurally animated gaze. *Proceedings of Motion on Games - MIG 13*. doi: 10.1145/2522628.2522630

Otta, E., Lira, B. B. P., Delevati, N. M., Cesar, O. P., & Pires, C. S. G. (1994). The effect of smiling and of head tilting on person perception. *The Journal of Psychology*, *128*(3), 323–331. doi: 10.1080/00223980.1994.9712736

Rossano, F. (2012). Gaze in conversation. *The Handbook of Conversation Analysis*, 308–329. doi: 10.1002/9781118325001.ch15

Ruhland, K., Peters, C. E., Andrist, S., Badler, J. B., Badler, N. I., Gleicher, M., ... Mcdonnell, R. (2015). A review of eye gaze in virtual agents, social robotics and hci: Behaviour generation, user interaction and perception. *Computer Graphics Forum*, *34*(6), 299–326. doi: 10.1111/cgf.12603

Scassellati, B. (1996). Mechanisms of shared attention for a humanoid robot. *In Proceedings of the AAAI Fall Symposium on Embodied Cognition and Action*, *4*, 21.

Sidner, C. L., Lee, C., Morency, L.-P., & Forlines, C. (2006). The effect of head-nod recognition in human-robot conversation. *Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction - HRI 06*. doi: 10.1145/1121241.1121291

Stanton, C., & Stevens, C. J. (2014). Robot pressure: The impact of robot eye gaze and lifelike bodily movements upon decision-making and trust. *Social Robotics Lecture Notes in Computer Science*, 330–339. doi: 10.1007/978-3-319-11973-1_34

Strabala, K. W., Lee, M. K., Dragan, A. D., Forlizzi, J. L., Srinivasa, S., Cakmak, M.,

& Micelli, V. (2013). Towards seamless human-robot handovers. *Journal of Human-Robot Interaction*, *2*(1), 112–132. doi: 10.5898/jhri.2.1.strabala

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257–285.

Szafir, D., & Mutlu, B. (2012). Pay attention! *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI 12*. doi: 10.1145/2207676.2207679

Team, T. w. (2019, Dec). *Welcome to wxpython!* Retrieved from `https://www.wxpython.org/`

Terzioğlu, Y., Mutlu, B., & Şahin, E. (2020). Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration. In *Proceedings of the 2020 acm/ieee international conference on human-robot interaction* (pp. 343–357).

Vertegaal, R., Slagter, R., Veer, G. V. D., & Nijholt, A. (2001). Eye gaze patterns in conversations. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI 01*. doi: 10.1145/365024.365119

Video conferencing, web conferencing, webinars, screen sharing. (n.d.). Retrieved from `https://zoom.us/`

Zaraki, A., Mazzei, D., Giuliani, M., & Rossi, D. D. (2014). Designing and evaluating a social gaze-control system for a humanoid robot. *IEEE Transactions on Human-Machine Systems*, *44*(2), 157–168. doi: 10.1109/thms.2014.2303083

APPENDIX A

QUESTIONNAIRES

A.1 Godspeed Questionnaire

Rate below questions in between 1 to 10.

1. Please rate your impression of the robot on these scales - Fake:Natural
2. Please rate your impression of the robot on these scales - Machinelike:Humanlike
3. Please rate your impression of the robot on these scales - Unconscious:Conscious
4. Please rate your impression of the robot on these scales - Artificial:Lifelike
5. Please rate your impression of the robot on these scales - Moving Rigidly:Moving Elegantly
6. Please rate your impression of the robot on these scales - Dead:Alive
7. Please rate your impression of the robot on these scales - Stagnant:Lively
8. Please rate your impression of the robot on these scales - Mechanical:Organic
9. Please rate your impression of the robot on these scales - Inert:Interactive
10. Please rate your impression of the robot on these scales - Apathetic:Responsive
11. Please rate your impression of the robot on these scales - Dislike:Like
12. Please rate your impression of the robot on these scales - Unfriendly:Friendly
13. Please rate your impression of the robot on these scales - Unkind:Kind
14. Please rate your impression of the robot on these scales - Unpleasant:Pleasant
15. Please rate your impression of the robot on these scales - Awful:Nice
16. Please rate your impression of the robot on these scales - Incompetent:Competent
17. Please rate your impression of the robot on these scales – Ignorant:Knowledgeable
18. Please rate your impression of the robot on these scales - Irresponsible:Responsible
19. Please rate your impression of the robot on these scales - Unintelligent:Intelligent

20. Please rate your impression of the robot on these scales - Foolish:Sensible
21. Please rate your emotional state while working with the robot on these scales - Anxious:Relaxed
22. Please rate your emotional state while working with the robot on these scales - Calm:Agitated
23. Please rate your emotional state while working with the robot on these scales - Quiescent:Surprised

A.2 Toolkit for Measuring Acceptance of an Assistive Social Robot

Rate below questions in between (Totally Agree/ Agree/ Not Sure / Disagree / Totally Disagree).

1. I would trust the robot if it gave me advice:
2. I would follow the advice the robot gives me
3. If I should use the robot, I would be afraid to make mistakes with it:
4. If I should use the robot, I would be afraid to break something with it:
5. I find the robot scary:
6. I find the robot intimidating:
7. I think it is a good idea to use the robot:
8. The robot would make my life more interesting:
9. It's good to make use of the robot:
10. I would like to work with the robot again:
11. I would like to work with the robot at work on a daily basis:
12. I think the robot can be adaptive to what I need:
13. I think the robot will only do what I need at a particular moment:
14. I think the robot will help me when I consider it to be necessary:
15. I enjoy the robot interacting with me:
16. I enjoy doing things with the robot:
17. I find the robot enjoyable:
18. I find the robot fascinating:
19. I find the robot boring:
20. I consider the robot a pleasant working partner:

21. I find the robot pleasant to interact with:
22. I feel the robot understands me:
23. I think the robot is nice:
24. I think the robot is useful to me "
25. It would be convenient for me to have the robot:
26. I think the robot can help me with many things:
27. When interacting with the robot I felt like I'm working with a real animal:
28. It sometimes felt as if the robot was really looking at me:
29. I can imagine the robot to be a living creature:
30. I often think the robot has no intelligence:
31. Sometimes the robot seems to have real feelings:
32. I would trust the robot to work close to me without harming me:
33. I would not be worried to be working in close proximity to the robot:

A.3 TIPI Questinnaire

Rate below questions in between (Totally Agree/ Agree/ Not Sure / Disagree / Totally Disagree).

1. I see myself as an introvert
2. I see myself as a reliable person
3. I see myself as a lazy person
4. I see myself as a comfortable person that handles stress successfully
5. I see myself as a person with an interest in art
6. I see myself as a social, outgoing person
7. I see myself inclined to find other people's mistakes
8. I see myself as someone who does their job carefully
9. I see myself as a person who gets angry quickly
10. I think that I have an active imagination