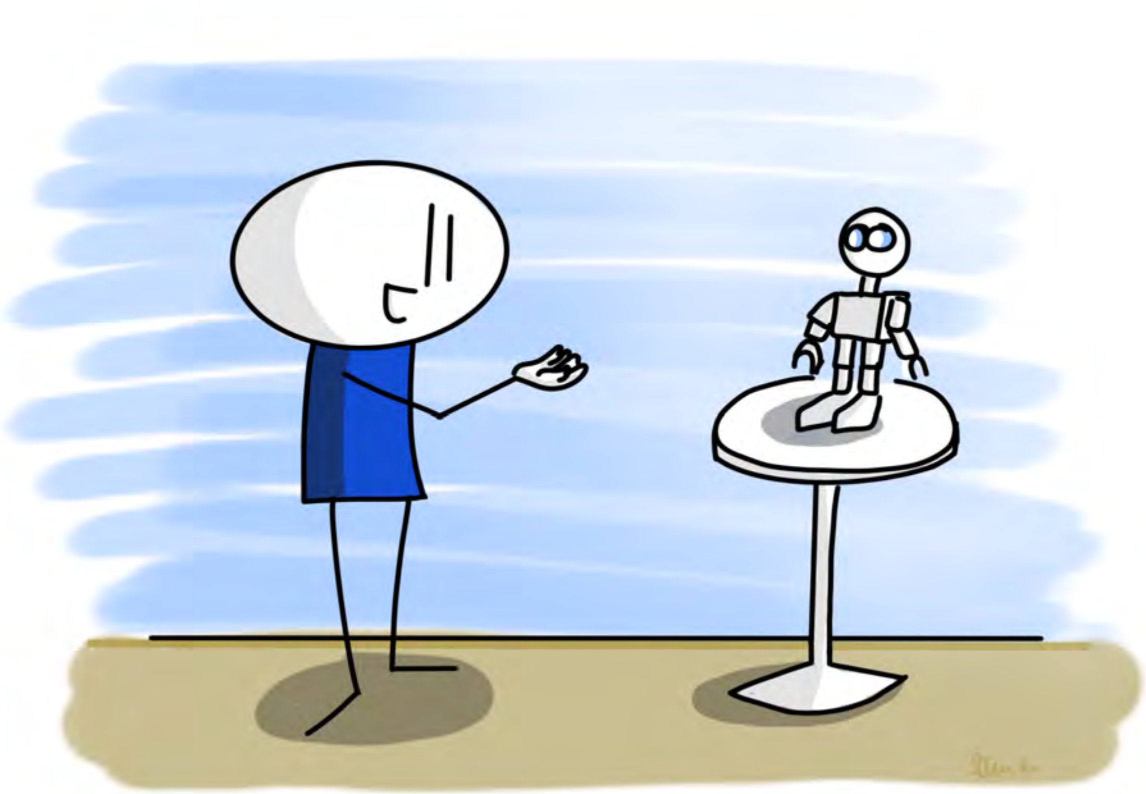
University of Southern Denmark
PhD Thesis

Effects of Contingent Robot Response to the Situatedness of Human-Robot Interactions



Author:
Lars Christian Jensen

Supervisor:
Kerstin Fischer

A thesis submitted in fulfillment of the requirements for the degree of PhD

in the

Graduate School at the Faculty of Humanities
Department of Design & Communication

SDU
UNIVERSITY OF
SOUTHERN DENMARK

August 31st, 2018

Printed by: Print & Sign, SDU

Summary

This thesis investigates how people respond to robots displaying situation awareness of certain contextual features in interaction. Displays of awareness of context, and how they are responded to are crucial in grounding a joint understanding of a situation, and building common ground between partners in interactions. I investigate how robots can signal to people how they understand a situation, and how these signals are understood by people. Situation awareness in Human-Robot Interaction usually considered a problem of engineering, where focus lies on building bigger and better sensors. However, in this thesis I treat situation awareness as a communication problem. As such, I systematically investigate the effects of particular *displays* of situation awareness, rather than situation awareness itself. Thus, the overarching research question that guides this investigation is:

- **What are the effects of a robot's displays of awareness to context?**

Analyzing displays allows prediction of how current and future sensory technologies can affect human-robot interaction, and shows specifically how a robot contributes to the joint understanding of common ground.

Theory and Methods

Chapter 1 introduces motivation for the research and introduces the theoretical frame through which the research question is explored. The investigation is informed by studies of interaction between people, Conversation Analysis, and studies of interaction between people and robots, Human-Robot Interaction. In Chapter 1 I show how I understand *context*, *situation awareness* and *common ground* to be related. I furthermore argue how common ground, as understood by Clark (1996), is a useful theoretical frame through which to evaluate human-robot interactions.

Chapter 2 introduces the methods used to design, execute and analyze the empirical investigations. Analytically, the thesis relies on two strikingly different methodologies; inferential statistics and Conversation Analysis. In Chapter 2 I argue for the use of each of these methodologies and account for why I believe that the combination of the two contributes more to our understanding of the topic, than either could have done on its own.

Empirical Investigation

The empirical investigation begins with Chapter 3. This chapter explores several aspects of *timing* and *contingency* as signals for common ground. In particular it attempts to address

the question whether contingent gaze successfully performs or contributes to certain social actions because it is contingent, or because the combination of gaze and contingency creates a unique social signal. This question is investigated through two research questions:

- How do contingent non-verbal responses affect how participants respond, adjust to, and perceive the robot in comparison to random responses?
- How do participants respond, adjust to and perceive the robot differently in the two contingent conditions?

In order to address these questions I designed a between-subject experimental study, in which participants tutor a small humanoid robot on English sentence construction. The study is designed with three conditions, *contingent gaze*, *contingent nods*, and *random gaze*. In the *contingent gaze* condition, the robot follows the gaze direction of the participant, so that when a participant gazes towards a certain object, so does the robot. In the *contingent nod* condition the robot's gaze is static but responds to the participants' verbal and nonverbal actions by nodding. In the *random gaze* condition, the robot's gaze is random. That is, the robot does not respond to the participant's gaze behavior.

This study is analytically the most diverse of the studies presented in the thesis, with analyses of participants' self-rated perception of the robot, as well as analyses of participants' gaze behavior and linguistic production. The study shows how contingent gaze contributes to broader joint understanding of the common ground and how this affects the ensuing interaction.

Chapter 4 investigates a different aspect of common ground, namely awareness of what has occurred in the interaction already. Specifically, the study asks the question:

- What are the effects on perception of a robot displaying an attention to previous events in the interaction?

Displaying an awareness of or an attention to what has already been said and done signals that local interactional history, the *discourse record*, can be considered to be common ground. This question is explored in a between-subject experiment in which a small humanoid robot instructs a human participant to construct a Lego figure. The experiment features two conditions, called *low aware* and *high aware*. In the *high aware* condition the robot make specific references to the perceptual basis, and to the discourse record. During the introduction of the experiment the robot asks participants whether he or she likes to play with Lego. During the end of the interaction it recalls the participants' response. The robot then asks participants if he or she thinks this activity was fun (in case they said they do like to play with Legos) or if it was fun despite their previous negative stance toward Lego. The robot also comments on the state of the weather. While the robot in the *low aware* condition asks participants whether they like playing with Lego, it never recalls the response.

The study finds that the robot is perceived to be more aware, more social, and more interactive with awareness manipulations than without.

Chapter 5 studies several aspects of common ground. The chapter investigates how displays of awareness to the *perceptual basis*, *face-tracking*, and *incremental* feedback each contributes to a joint understanding of the common ground. These three types of signals of common ground are investigated in an experimental study in order to better understand how they affect perception and behavior for participants in interaction with a robot. The aim of this research is to understand the relative contributions of each of these signals to participants' perception and behavior. In the experiment, a small humanoid robot guides participants through a series of physical exercises. Participants' perception of the robot is evaluated in a post-experiment questionnaire, while their behavior is evaluated by how much water they drink during the exercise, and the extent to which they follow the robot's prompt to drink water.

The study shows that each of the three displays to contextual information contributes in different ways to participants' perceptions and to the interactions themselves. There is very little overlap between conditions, which serves to show that each of the displays contribute differently. That is to say, the kind of contextual information a robot displays an awareness of has a large impact on how it is perceived and responded to.

Chapter 6 investigates perceptual and interactional effects of incremental feedback. Specifically, I investigate:

- What are the perceptual and behavioral effects of incremental feedback?

In addition, I also explore how behavior relates to perception. These questions are investigated in an experimental study in which a mobile robot guides a human participant around in an office space to collect certain items. The experiment is carried out in a between subject experimental design with two experimental conditions. In one condition the robot is able to modify its speech incrementally based participants non-verbal conduct. On two occasions, as participants are looking for certain items, the robot can direct their search by producing utterances like "more to the right" and "yes a little more". In the other condition the robot says approximately where the object can be found, but offers no additional advice.

The study finds that incremental feedback enables participants to perform better. Adding incremental feedback to a robot's communication design increases the perceived common ground between robot and participants. This also means that when participants perform poorly, they hold the robot responsible, as evidenced by lower ratings in those cases.

Chapter 7 compares the perceptual and performative effects of two gaze behaviors, proactive and reactive gaze, in a collaborative assembly scenario. The focus of the chapter is to investigate how displays of contextual awareness through contingent robot responses affect interaction and perception. This is explored in a controlled experiment with an industrial robotic platform.

In the experiment, participants are asked to assemble an IKEA children's stool with the assistance of the robot. Their task is to instruct the robot to fetch the legs of the stool, while the participants themselves have to perform the actual assembly. It is left open to the participants exactly how to instruct the robot. The instruction consists of two phases: a fetching phase, in which participants have to indicate to the robot which of the four legs they want, and a handover phase, in which the participant have to let the robot know where to deliver the leg. The participants then connect the leg to the seat until all four legs are in their respective slots, and the chair is assembled.

The experiment has two conditions in a between-subjects design. Initially, the robot looks at, and tracks participants' faces until the robot starts moving its arm. In one condition, the robot gazes proactively. That is, whenever the robot arm moves from one location to another, the robot head indicates where it moved to, by gazing to this location in the workspace prior to and during robot arm movement. Both, head pose and eyes fixate on the target location. In the other condition, the robot gazes reactively. That is, whenever the robot arm moves, the robot head 'follows' the arm via a tracking motion. This is referred to as the reactive condition. After each move, the robot face returns to look at, and track the face of the participant, until it receives a new instruction.

Participants' perception of the robot is evaluated from a post-experiment questionnaire, while their behavior is evaluated in terms of their gaze and pointing behavior. The study shows that participants do not evaluate the robot's gaze as a signal of an understanding of a joint plan. Thus, proactive gaze cannot be shown to signal an understanding of the common ground.

Chapter 8 is the final empirical chapter of the thesis. In it I present a study of an experiment set in an identical setup as the experiment presented in Chapter 8. The chapter investigates how a robot is able to display its awareness towards certain aspects of participants' communication with it, by responding to repair initiated by participants, after the robot has made an error. The error made by the robot is not fatal, or even critical, but is treated by participant as interaction trouble. More specifically, this chapter investigates a display of contextual awareness, in which a robot is able to change its online behavior, based on a human communication partner's gestural action.

The experiment has two conditions in a between-subjects design. In one condition, the robot is able to change its current actions based on participants' gestural activity. In other words, the robot is able to, in real time, respond to participants' repair initiations. In the other condition the robot responds only to the first instruction given by participants, and is thus not able to respond to repair initiations. The study shows that implementing just one opportunity for repair, can significantly affect participants' perception and behavior. Specifically, results show that the response to repair updates participants' partner model of the robot, and subsequently changes how they interact with the robot, which methods they use, and how they perceive the robot.

Implications

The last chapter discusses findings from each of the empirical chapters and relates them to the conceptual model of common ground introduced in Chapter 1. Specifically, four out the five indicators for situation awareness are found to contribute to common ground. Furthermore, results show that the more a robot can display its awareness to context, the more favorable it is perceived, and the more seriously it is treated as an interaction partner. However, there is a caveat. The more situationally aware a robot displays to be, the more users expect it to be able to perceive, understand and do. This may cause users to overestimate its abilities, which can have problematic consequences for the interaction.

Finally, I discuss how the results obtained in the thesis might inform design decisions for future robots.

Resumé

Denne afhandling undersøger hvordan mennesker forholder sig til robotter der viser tegn på situationsfornemmelse. Hvordan mennesker forholder sig til tegn på situationsfornemmelse i interaktioner er yderst vigtigt for at skabe forståelse mellem interaktionspartnere. Jeg undersøger hvordan robotter kan signalere til mennesker hvordan de (robotterne) forstår en given situation, og hvordan mennesker forstår og tolker sådanne signaler. Situationsfornemmelse i menneske-robot interaktion er traditionelt set et ingienørproblem, hvor fokus ligger på bedre og større sensorer. I denne afhandling behandler jeg det dog som et kommunikationsproblem. Som et kommunikationsproblem undersøger jeg systematisk effekterne af specifikke tegn på situationsfornemmelse. Det overordnede forskningsspørgsmål er således:

- **Hvad er effekterne af en robots tegn på situationsfornemmelse?**

Analyse af forskellige tegn på situationsfornemmelse kan give et fingerpeg om hvordan nutidige og fremtidige teknologier kan indøve inflydelse på menneske-robot interaktion. Mere specifikt kan en sådan analyse også bidrage til en bedre forståelse for hvordan fælles forståelse mellem interaktionspartnere opstår og vedligeholdes.

Teori og Metode

Kapitel 1 introducerer motivationen for den forskning der præsenteres i afhandlingen og introducerer også den teoretiske ramme gennem hvilken forskningsspørgsmålet bliver undersøgt. Undersøgelsen der er foretaget i afhandlingen er bygget på studier af interaktion mellem mennesker (konversationsanalyse) og studier af interaktion mellem mennesker og robotter (menneske-robot interaktion). I dette første kapitel introducerer jeg koncepterne, forklarer hvad *kontekst*, *situationsfornemmelse* og *fælles forståelse* er og forklarer hvordan de er forbundet. Derudover redegør jeg for hvorfor jeg mener at fælles forståelse, som forstået af Clark (1996) er en nyttig ramme gennem hvilken man kan evaluere menneske-robot interaktioner.

Kapitel 2 introducer de metoder der er anvendt til at designe, udføre og analysere de empiriske undersøgelser. Afhandlingen beror analytisk på to meget forskellige metodologier; statistisk metode og etnometodologisk konversationsanalyse. I kapitlet argumenterer jeg for brugen af hver af disse to metodologier og redegør for hvorfor jeg mener at kombinationen af disse to bidrager til mere en hvad hver af dem ville kunne bidrage hver for sig.

Empiriske Undersøgelser

Den empiriske undersøgelse begynder i kapitel 3. Dette kapitel udforsker adskillige aspekter af timing og responsivitet som signaler for fælles forståelse. Mere præcist, forsøger jeg med kapitlet at adressere spørgsmålet hvorvidt responsiv synsretning bidrager til social interaktion fordi det netop er responsivt, eller om det er kombinationen af synsretning og responsivitet der sammen sender et unikt socialt signal. Dette spørsmål er undersøgt gennem to forskningsspørgsmål:

- Hvilken indflydelse har non-verbale udtryk på hvordan deltagere forholder sig til, tilpasser sig og opfatter en robot i forhold til hvis robotten anvendte tilfældige non-verbale udtryk?
- Hvad er forskellene på hvordan deltagere forholder sig, tilpasser sig og opfatter robotter der bruger en af to forskellige måder at udvise responsivitet på?

For at kunne adressere disse spørgsmål har jeg udfærdiget et eksperiment, i hvilket deltagere underviser en lille humanoid robot in engelsk sætningskonstruktion. Studiet har tre scenarier. Ét scenarie hvor robottens synsretning er responsiv, ét scenarie hvor robotten nikker responsivt, og ét scenarie hvor robottens synsretning er tilfældig. I det første scenarie følger robottens synsretning hele tiden deltagerens synsretning, så når en deltager ser hen imod et bestemt objekt, ser robotten også i den retning. I det andet scenarie ser robotten altid kun i én retning, men udviser responsivitet ved at nikke efter deltageres talehandlinger. I det sidste scenarie kigger robotten tilfældigt rundt i rummet og er helt uafhængig af hvad deltageren laver eller siger.

Dette studie er analytisk set det mest mangfoldige blandt de studier der er i afhandlingen. Studiet analyserer deltageres spørgeskemabesvarelser, synsretning, og sproglig produktion. Studiet viser hvordan robottens responsive synsretning bidrager til en bredere fælles forståelse, og hvordan denne fælles forståelse har indflydelse på interaktionen.

Kapitel 4 undersøger et anderledes aspekt af fælles forståelse, mere præcist forståelse for den lokale interaktionshistorik. Studiet forsøger at svare på spørgsmålet:

- Hvilken indflydelse har det på deltageres opfattelse af en robot, at den er i stand til at udvise en forståelse for handlinger er der foretaget tidligere i en interaktion?

Tegn på opmærksomhed for hvad der er allerede er blevet sagt og gjort i en interaktion, signalerer at den lokale interaktionshistorik kan betragtes som en del af den fælles forståelse mellem interaktionspartnere. Dette spørgsmål er undersøgt i et kontrolleret eksperiment, i hvilket en lille humanoid robot instruerer en menneskelig deltager i at bygge en Lego model. Eksperimentet har to scenarier, der henvises til som *lav opmærksomhed* og *høj opmærksomhed*. I scenariet med *høj opmærksomhed* laver robotten særlige henvisninger til den lokale interaktionshistorik. I starten af eksperimentet spørger robotten deltageren om denne kan lide at lege med Lego. Igen imod slutningen af eksperimentet siger robotten hvad deltageren havde svaret og spørger dertil om deltageren synes at denne aktivitet havde været sjov (i tilfælde at deltageren godt kunne lide at lege med Lego) eller om aktivitet var sjov *på trods af* at deltageren ikke kunne lide at lege med Lego (i tilfælde hvor deltageren sagde at denne ikke kunne lide at lege med Lego). Derudover kommenterede robotten også på vejret (hvorvidt vejret var godt eller dårligt) i scenariet med *høj opmærksomhed*. Robotten med *lav opmærksomhed* spurgte også om deltagere kunne lide at lege med Lego, men anvendte ikke svaret senere i interaktionen.

Studiet viser at robotten blev opfattet som mere opmærksom, mere social og mere social interaktiv i scenariet med *høj opmærksomhed* end uden.

Kapitel 5 undersøger tre forskellige aspekter af fælles forståelse. Kapitlet undersøger hvordan tegn på opmærksomhed på begivenheder, ansigtssporing og trinvise feedback bidrager til en fælles forståelse af en interaktionssituation. Disse tre tegn på fælles forståelse bliver undersøgt i et kontrolleret studie for bedre at forstå hvordan de bidrager til hvordan deltagere forholder sig til og opfatter en robot. Formålet er at forstå hvordan hver af disse tegn bidrager til en fælles forståelse mellem menneske og robot. I eksperimentet guider en lille humanoid robot deltagere igennem en række fysiske øvelser. Deltageres opfattelse bliver evalueret gennem en spørgeskemaundersøgelse, og hvordan de forholder sig til robotten bliver evalueret ved at måle hvor meget vand de drikker under eksperimentet og hvorvidt de følger robottens opfordringer om at drikke vand.

Studiet viser at hver af disse tre tegn bidrager på forskellige måder til deltagernes opfattelse af robotten og selve interaktionen. Der er ganske lidt overlap mellem de tre scenarier, der viser at hvert tegn bidrager til forståelsen på forskellige måder. Det vil sige at, afhængig af hvilke tegn på opmærsomhed robotten udviser, opfatter deltagere robotten anderledes og handler ligeledes anderledes.

I kapitel 6 undersøger jeg *trinvis feedback* lidt nærmere. Mere præcist undersøger jeg:

- Hvordan bidrager trinvis feedback til deltageres opfattelse af robotten og hvordan de forholder sig til den?

Derudover undersøger jeg også hvordan deltageres handlinger relaterer til deres rapporterede opfattelser. Disse spørgsål er undersøgt i et kontrolleret eksperiment hvor en mobil robot guider en menneskelig deltager rundt i et laboratorie for at indsamle en række genstande. Eksperimentet er udført med to scenarier. I et scenarie er robotten i stand til at trinvist ændre dens talehandlinger, hvilket den gør med basis i deltageres non-verbale handlinger. På to forskellige tidspunkter i eksperimentet guider robotten således deltageren til hvordan de kan finde den genstand robotten har bedt dem om at finde. Dette gøres ved at robotten f.eks. siger "du skal lidt mere til højre" og "ja en lille smule mere". I det andet scenarie giver robotten kun en beskrivelse af hvor genstanden cirka kan findes.

Studiet viser at trinvis feedback gør deltagere i stand til finde objekterne hurtigere. Det vil sige at ved at tilføje trinvis feedback til robottens kommunikationsdesign kan man øge den fælles forståelse mellem robot og menneske. Dette betyder dog også, at når deltagerne har problemer med at finde objekterne giver de robotten skylden, hvilket kan ses i mere negative bedømmelser når dette sker.

Kapitel 7 undersøger effekterne af to forskellige måder at regulere synsretning på i en industriel robotplatform. I eksperimentet bliver deltagere bedt om at sammen med robotten samle en børnestol fra IKEA. Deltagernes opgave er at instruere robotten i at give dem de rigtige dele, og så selv samle stolen når de har fået delene. Deltagerne får ikke eksplicit at vide præcis hvordan de skulle instruere robotten. Instruktionen består af to faser; en ‘hente’ fase hvor deltagere skal indikere til robotten hvilken en af de fire dele de vil have robotten til at hente, og en ‘overdragelse’ fase hvor deltagere skal vise robotten hvor delen skal hen. Dette gentager de indtil stolen er samlet.

Eksperimentet har to scenarier. I et scenarie kigger robotten proaktivt hen til det område hvor den er på vej hen. Robotten udviser altså en forståelse for den fælles plan. I det andet scenarie følger robotten altid kun dens egen bevægelser. Det vi sige at, når robottens arm bevæger sig følger robottens hoved og øjne armen. Uanset scenarie, kigger robotten altid tilbage på den menneskelige deltager når den er klar til en ny kommando.

Deltagernes opfattelse af robotten er evalueret gennem en spørgeskemaundersøgelse, men hvordan de forholder sig til robotten er evalueret gennem en analyse af deres pegeadfærd. Studiet viser at deltagerne ser ikke robottens proaktive synsretning som et tegn på forståelse af en fælles plan. Altså kunne det ikke påvises at proaktiv synsretning bidrager til den fælles forståelse i interaktion.

Kapitel 8 præsenter det sidste empiriske studie i afhandlingen. I dette kapitel præsenter jeg et studie som der i dens opsætning er identisk med studiet i det forrige kapitel. Kapitlet undersøger hvordan en robot er i stand til at udvise en opmærksomhed for særlige aspekter at deltageres kommunikation med den, ved at være i stand til at kunne reagere på deltageres reparaturer, efter at robotten har lavet en fejl. Eksperimentet har to scenarier. I et scenarie er robotten i stand til at ændre dens adfærd på baggrund af deltageres pegeadfærd. Med andre ord er robotten i stand til reagere på deltageres reparaturer. I det andet scenarie reagerer robotten kun på deltageres første instruktion og ignorerer alle andre instruktioner. I dette scenarie kan robotten altså ikke reagere på reparaturer. Studiet viser at selv små muligheder for reparaturer kan resultere i væsentlige ændringer i hvordan deltagere opfatter og forholder sig til robotten. Mere specifikt viser studiet at deltagere i det første scenarie har langt større muligheder for at opdatere deres partnermodel, hvilket ændrer hvordan de instruerer robotten, hvordan de forholder sig til robotten og hvordan de opfatter robotten.

Konklusioner

Det sidste kapitel diskuterer resultater fra hver af de empiriske kapitler og relaterer dem til den konceptuelle model for *fælles forståelse*, introduceret i kapitel 1. Studierne viser at fire ud af de fem undersøgte tegn på situationsfornemmelse bidrager til en fælles forståelse. Derudover viser studierne at jo mere en robot kan udvise dens situationsfornemmelse og dens opmærksomhed til konteksten jo mere positivt bliver den bedømt og jo mere seriøst bliver den taget som en interaktionspartner. Der dog en modhage. Jo mere en robot udviser tegn på situationsfornemmelse, jo større forventninger har deltagere også til hvad robotten er i stand til at forstå og gøre. Dette kan skabe en situation hvor interaktionspartnere kan overvurdere en robots færdigheder, hvilket kan skabe problemer i interaktionen.

Endeligt diskuterer jeg hvordan de resultater jeg præsenterer kan bruges i desingbeslutninger af fremtidige robotsystemer.

Acknowledgments

Foremost, I want to thank my supervisor for her guidance and support. Kerstin has opened more doors for me than I can count, and continues to be a source of inspiration as a researcher. I am forever grateful for all the possibilities she has given me. I look forward to continued collaboration. However, for the record, I will just say that lunch at 11.30 is not a minute too early!

I also want to thank the research institutions that I have visited over the years. I had the pleasure of visiting Katrin Lohan and Ingo Keller at Heriott-Watt in Edinburgh, Scotland in the beginning of 2016. I want to thank especially Ingo for numerous tutorials on the inner workings of YARP. I also want to thank Hoang-Long Cao and Bram Vanderborght, whose lab I visited in Brussels, Belgium in the fall of 2016. Thanks also to Justus Piater, Dadhichi Shukla, Özgur Erkennt, and Sebastian Stabinger for making me feel at home when I visited the Intelligent and Interactive Systems lab at the University of Innsbruck, Austria. My many travels for research visits and conference participation would not have been possible without aid from my close friends and family to look out for my kids in my absence. I especially want to thank Maiken and Stina, Torben, Maria, and my mom.

Many people have also assisted me in running experiments and collecting data. Here I want to especially thank Yong Ding, Rosalyn, Anna, Nadine, Maria, Selina, Emanuelly, and Franziska.

I also want to thank Nathalie Schümchen for creating the illustrations that adorns the front page of the thesis.

Thanks also to the Research Network for Transdisciplinary Studies in Social Robotics (TRANSOR) and Johanna Seibt for many fruitful discussions and travel grants that allowed me to present my work at several international conferences. I also want to thank Fabrikant Mads Clausens Fond and the Torben & Alice Frimodt Fond for grants that has allowed our lab to purchase the robots, computers, sensors and recording equipment I have used for my thesis work. Part of the work I present in the thesis is also funded by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 610878, 3rdHAND.

The online community at Stack Exchange (stackexchange.com) have also been a huge help, so thank you for being there!

Finally, I want to especially thank my partner in life, crime and everything work-related, Maria, Mary, Mawily.

Table of Contents

1. Introduction

1.1 Problem Statement and Motivation

A huge problem for people interacting with robots is that they often do not know how robots perceive the world and the people and objects in it. This becomes a problem when people need to engage in joint interactions with robots. Without a shared basis for perception, interactions are prone to interactional trouble. The problems that can arise in interactions with technology are very well described by Suchman (2007). In her study of how people use a photocopying machine she showed how communicative breakdowns happen when humans and machines do not have access to the same kinds of information, and when they make false assumptions about what the other can see. The problem described by Suchman also holds for robots; people do not know how or what a robot perceives and do not know how to find out either. Evidence of the problem for Human-Robot Interaction (HRI) is found in accounts of trouble in human-robot interactions. For example, Jensen, Fischer, Suvei, and Bodenhagen (2017) report on the difficulties users have in understanding requests made by a robot, and Gehle, Pitsch, Dankert, and Wrede (2015) show that participants display confusion when a robot acts an unexpectedly.

People interacting with other people do not face the same problems to the same extent. In interactions with others, people can already make certain assumptions about their communication partners. People can reasonably expect that other humans have senses, such as vision, hearing or smell, that function in similar ways as their their own. This means that when people see a cup, for example, they expect that other people close to them also see a cup. Not only do people assume other people to also to see its shape and color, they also assume them to know know how to hold and use it. None of these assumptions necessarily hold for robots, and when people do make such assumptions they usually encounter trouble.

One way to circumvent problems because of differences in perception is by using a translation system, using augmented reality markers, such as QR-codes or hamming makers (perceivable and meaningful for robots), to represent specific objects that are perceivable and meaningful for people (Huang & Mutlu, 2016; Mihalyi, Pathak, Vaskevicius, Fromm, & Birk, 2015). However, this and other similar methods do not ground understanding between robots and people, but rather create a bridge between two ways of perceiving and understanding the world (Searle, 1980). Furthermore, as robots are expected to engage in increasingly complex social situations, robots will be expected to perceive and understand not only simple objects, but also concepts, relations and social cues, which may not be as easily 'translated' or bridged.

The problem of perception is most often treated as an engineering problem, solved with more and better sensors and new machine learning techniques. While these technological advances definitely change what robots can do and how people think about them, people have no better understanding of how robots perceive and understand the world than Suchman's users had of their photocopying machine (2007). While the problem most often is treated as one of engineering, it may also be useful to consider it as a problem of communication. One aspect of the problem of diverging perceptions and understandings of the world is that the knowledge that people and robots hold is not grounded in a joint understanding of the situation they are in. Grounding is a process in which participants in interaction update their understanding of the common ground between them on moment-by-moment basis (Clark & Brennan, 1991). In other words, people in interaction make observable to each other what aspects of an interaction they consider to be jointly understood by all parties involved. Treating the problem of perception and understanding in human-robot interaction as a problem of (lack of) common ground has implications for how the problem can be addressed. Specifically, the number of and complexity of the sensors a robot has moves to the background, while the question is how robots can signal how and what it perceives and understands gains more importance. How this signaling can be achieved and what it means for interaction between robots and people is what this thesis explores.

More specifically, I theoretically and empirically investigate how robots' displays of awareness of participants, their behavior, and the context in which the interaction takes place affects interaction and how people perceive robots.

The aim with this thesis is to find out how people display situation awareness to certain contextual features in social interaction.

1.1.1 Research Question

I now turn to exactly what will come under investigation. The central research question of the dissertation is, **what are the effects of a robot's displays of awareness to context?** Displays (a term that is borrowed from the conversation analytical terminology) refer to the practices that communication partners make in order to make resources and understandings visible to each other. The focus of displays in conversation analytical work stresses the point that relevance of practically anything that goes on during interaction is negotiated by communication partners in the way they respond to it. The implication for the research question here is that a robot needs to make visible to its human communication partner whether it is aware of situational aspects of an interaction. Thus, displays work as an indicator for the common ground (Clark, 1996, p. 95). The responses to these indicators are what is under investigation. The research question is explored systematically through six empirical studies in which several verbal and nonverbal indicators for situational awareness are implemented in three different robotic systems.

1.2 Theoretical Framework

In this section, I describe the theoretical framework that I draw upon to guide my investigations.

1.2.1 What is Context?

Context is in layman's terms understood as the circumstances under which an event takes place. Context can be anything from where an event takes place, when it takes place, who participates, how the event come to be, etc. This is also what formal semanticists and positivist research paradigms understand by context (Kamp & Reyle, 1993; Kamp & Roßdeutscher, 1992). In this view, context is not negotiated but treated as unproblematic. Context in this understanding focuses on for example participants (e.g. doctors and patients), the environment (e.g. a clinic) and an activity (e.g. consulting). These characterizations are given meaning regardless of whether communication partners attend to them.

Ethnographic Understanding of Context

A different conceptualization of context can be found in ethnography for example. This understanding can be observed in sociolinguistic research approaches, as in, for example, the 'Ethnography of Speaking' (Hymes, 1964), which formalizes and categorizes communication according to a set of predefined characteristics. For example, Holmes (1989) distinguishes between men and women in the way they communicate politeness. Aoki (2000) considers 'family' and 'religion' as relevant contexts for a study of Mexican Americans in California, and Samy Alim (2007) considers ethnicity as a relevant context. In this understanding of context, what happens during interaction and the way communication partners behave are not considered to be part of the context, but rather a *product* of context.

Ethnography of speaking also considers knowledge representations that cannot be considered 'factual' in the same sense that a layman's understanding of context does. For example, the topic and purpose of the communication and social norms are all considered to be part of the context in an ethnography of speaking. The approach highlights ethnographic differences among communication partners, some of which are usually only noticed by participants themselves when they experience breakdowns in communication (Holmes, 2008, p. 366).

However, ethnographic context can also include information, for example knowledge about contextualization cues (Gumperz, 1982), which is disclosed in communication as it occurs. Contextualization cues may signal assumptions communication partners have of each other and the situation they are currently in. Contextualization cues include, for example, language choice (Gafaranga, 2007), prosody (Culpeper, 2011), lexical choice and facial expression (Holmes, 2008, pp. 374–375).

Ethnographic context is used much in the description and analysis of intercultural encounters, and especially in intercultural miscommunication. Thus, troubles in communication between different speech communities and other cultural entities are explained in terms of the culture(s) communication partners belong to and how they interpret, and draw inferences from, contextualization cues. However, assigning explanatory power to cultural affiliations in communicative breakdowns has come under some critique (see for example, Sarangi (1994) and Holliday (1999)). They offer a view of culture that is more abstract and considers social groupings (e.g. a specific classroom or a specific workplace) as the largest cultural entity. In this view of ethnography contextualization cues are described as part of the behavior of a certain social group, but context as such has no explanatory power.

In summary, an ethnographic understanding of context is made up of communication partners' cultural, ethnic and linguistic affiliations and of how people deploy and interpret signals that communicate these affiliations. These signals can be linguistic, para-linguistic, or expressed through non-verbal behavior.

Ethnomethodological Context

An ethnomethodological conversation analysis (CA) perspective of context is quite different from laymen's and ethnographic understandings of context. Here, context is not given, but it is rather a locally established interactional resource for communication partners. This means that context does not predefine interaction or its interaction partners. In CA, context is not a pre-established feature of interaction, but it is whatever communication partners evoke during interaction (Schegloff, 1997). That is, context needs to be made relevant by communication partners themselves, and the understandings they bring to bear to the interactions needs to be observable. In CA, all features of an interaction can be considered as part of the context. However, especially the structural components of interaction, such as adjacency pairs (Sacks, Schegloff, & Jefferson, 1974), conditional relevance (Schegloff, 1968), timing (Jefferson, 1989), projections (Schegloff, 1980), turn-taking, prior utterances, and the next-turn proof procedure, are seen as relevant context. In CA, interaction is understood in the light of what has come before, what is projected to come next, and what else is happening in the immediate interaction space. In this sense, interaction is contextually situated. This means that specific utterances or actions are not taken to mean anything by themselves, but need to be negotiated and ratified by communication partners; therefore, people need to signal to each other continually what they understand the context to be.

In principle, what is considered as context in laymen's terms and in ethnographic approaches can also be considered as context in an ethnomethodological perspective. However, this approach comes with the caveat that these notions of context can only become relevant when participants in interaction display an orientation to them (and make them relevant). This can be done, for example, through membership categorization (Sacks, 1989). Sacks (1989, p. 273) notes that:

> "*If we're going to describe Members' activities, and the way they produce activities and see activities and organize their knowledge about them, then we're going to have to find out how they go about choosing among the available sets of categories for grasping some event.*"

He goes on:

> *"If any Member hears another categorize someone else or themselves on one of these items, then the way the Member hearing this decides what category is appropriate, is by themselves categorizing the categorizer according to the same set of categories."* (Sacks, 1989, p. 277)

Sacks stresses here that notions of context, in the form of participant characterization, are not given, but are revealed through participants' conduct. Correspondingly, Schegloff has expressed concerns about using objective and ethnographic notions of context as explanatory factors in accounts of social interaction on several occasions (Schegloff, 1987b; 1997). In particular he says that:

> *"It is being proposed that the much invoked "dependence, on context" must be investigated by showing that, and how, participants analyze context and use the product of their analysis in producing their interaction."* (Schegloff, 1972)

Schegloff's concerns are based on the central notion in CA that no feature of interaction or its participants can be taken for granted unless participants in interaction make observable that they orient to such feature. From this perspective, HRI can only be successful if a robot signals what it takes the context to consist of.

Context is this sense is very broad and can encompass many different types of observations. Therein lies the strength of an ethnomethodological understanding of context. However, in order to understand how context becomes relevant, it is necessary to look at how communication partners signal their attention to context. One way to do this is to look at such signals as indicators for common ground. That is, communication partners signal to each other that they take the common ground to be by attending to certain aspects of the context.

1.2.2 What is Common Ground?

The most complete account of common ground is given by Clark (1996), and revolves around a model of interaction in which people tailor their contribution in interaction to what they think they know their communication partners to know and to be interested in. Clark defines common ground as:

> *"...the sum of ... mutual, common, or joint knowledge, beliefs, and suppositions."* (Clark, 1996, p. 93)

That is, common ground is the information people take for granted or assume their interaction partner to know. Clark posits that common ground is achieved through an awareness towards who the interaction partner is and includes ethnographic information such as what their profession is, where they are from, what their hobbies or interests are, background information about the interaction, such as, where and when the interaction takes place, who is present, why they are there, and interactional information such as what has gone on in the interaction already, what is currently going on, and what is projected to come next. However, exactly which pieces of joint knowledge communication partners draw inferences from has implications for what they consider common ground. Thus Clark (1996, p. 99) argues that:

> *"When it comes to coordinating on a joint action, people cannot rely on just any information they have about each other. They must establish just the right piece of common ground, and that depends on the them finding a shared basis for that piece"*

People in interaction establish common ground through two shared resources; communal common ground and personal common ground (Clark, 1996, p. 100). Communal common ground is made up of information much of which can be described as information that is ethnographic in nature. This includes information about gender, ethnicity, occupation, and nationality (Clark, 1996, p. 103). Information of this kind allows people to make inferences about what their communication partners know and what they might be interested in. Communal common ground thus provides people with information that allows them to expand and solidify the assumptions people have of each other, which in turn enables joint action. Communal common ground consists of five elements; human nature, lexicons, cultural facts, ineffable background and the grading of information, and thus extends well beyond ethnographic information.

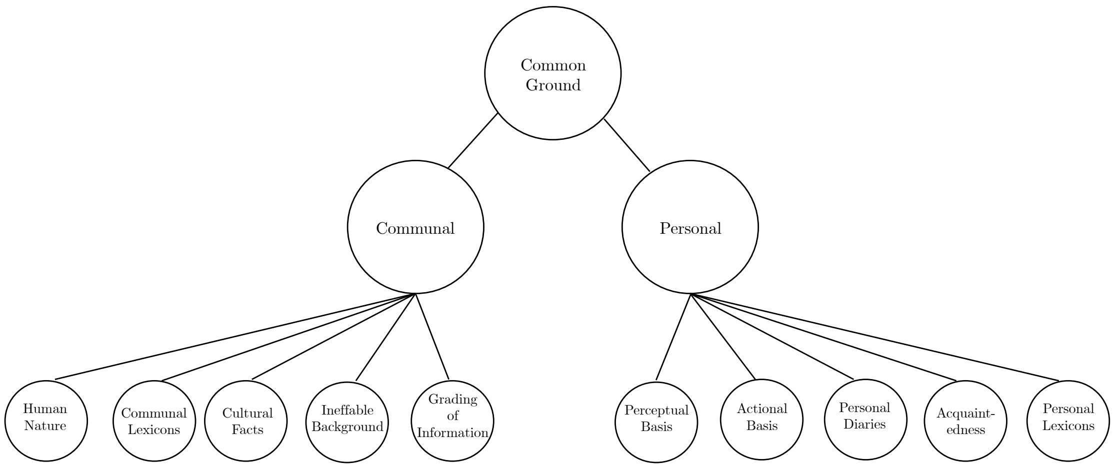


Figure 1.1: Common Ground

Communal Common Ground

Human nature, which is one of the five elements of common ground, refers to the assumptions people make about other people only from the fact that they indeed are human. For example, when meeting others, people make the assumption that other people possess the same senses, such as hearing, smell, and vision, and that these senses function in similar ways as their own. Although these assumptions might not turn out to be correct, according to Clark (1996, p. 106), they form the starting point from which to build common ground. Communal lexicons refer to the linguistic practices and special terminologies of social groups. For example, people who belong to the same profession, people who share a native language, or people who belong to the same neighborhood are assumed to share specialized linguistic knowledge that people outside those communities do not.

Cultural facts refer to the ethnographic knowledge people assume other people to have, based on the social groups they belong to or the geographic regions they come from, for example. Ethnographic knowledge includes the cultural facts and norms also covered in the objective and ethnographic context. Ineffable background includes the feelings associated with cultural facts.

Ineffable background can be summarized as facts that need to be experienced before they can be 'known'. For example, one can read about cycling or skiing, but will not know how it is to ski down a mountain, or drive through traffic on a bicycle before experiencing it.

The final element in communal common ground, the grading of information, refers to the ability people have to estimate what or how much other people may know.

Communal common ground allows people to draw inferences based in their own experience and knowledge, and the assumptions they have about what their communication partners have experienced, what they know, and what people think they know (Clark, Schreuder, & Buttrick, 1983).

Personal Common Ground

Another aspect of common ground is what Clark (1996, p. 112) refers to as personal common ground. This aspect takes into account not only what is currently going on in the interaction, but also what has come before and how well interaction partners know each other. Thus, personal common ground is based on current and previous joint experiences with interaction partners. Personal common ground consists of five elements: perceptual basis, actional basis, personal diaries, acquaintedness[1], and personal lexicons.

The perceptual basis can be described as an awareness to what is going on in the immediate environment. As the name implies, it refers to elements that are perceivable, such as objects in the interaction space or particularly salient events.

The actional basis refers to the joint actions, for example the talk communication partners are in involved, or playing chess.

Personal diaries refer to memory representations of earlier actional and perceptual experiences, which form the basis for the current common ground, but refer also to the discourse record of the current interaction. That is, personal diaries comprise the actional and perceptual basis that has taken place already.

Acquaintedness is, simply put, the level of acquaintance communication partners have with each other. That is, the more acquainted communication partners are, the more common ground they are assumed to have, as they would have shared more actional and perceptual experiences.

Finally, personal lexicons are an indicator for common ground that is expressed directly in the language communication partners use. Personal lexicons differ from cultural lexicons in that personal lexicons are not defined by members in certain social groups or communities,

[1] Clark, refers to this as 'friends and strangers'.

but are rather based on personal acquaintance. For example, lovers give each other nick names, and soldiers in a military unit give each others nick names, which only they share.

Personal common ground allows people to draw inferences from interaction as it happens and from previous interactions with the same people. These inferences enable communication partners to make assumptions about the shared common ground, which may have huge impact on how they interact. In interaction, people draw on both communal and personal common ground, using all the ten elements to establish and continuously update the shared common ground.

So far I have positioned context as the content matter and common ground as the mechanism through which partners in interaction select what aspects of context they attend to. In order for communication partners to make this selection, they need to be aware of the aspects of context that may be relevant. For humans this is not so problematic. As discussed above, there are aspects of the communal common ground, such at the human nature, that allow people to take a lot of things for granted. People interacting with robots cannot take the same things for granted, even though they sometimes do. Thus, in order for robots to successfully signal what they take the common ground to be, they must signal information about their awareness of the situation.

1.2.3 What is Situation Awareness?

Situation awareness (SA) is a concept that covers the perception of elements of events as they unfold, the comprehension of their meaning and salience, and a projection of what comes next (Endsley, 1995). Specifically, Endsley (1988) defines SA as:

> *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"*

Thus, according to Endsley, SA refers to the ability to find out what is going on in the immediate environment and what influences the actions people undertake. Dominguez, Vidulich, Vogel, and McMillan (1994) build on Endsley's work on SA, but also include psychological concepts, such as mental models, in their definition of context. According to them, people store information from their environments in mental models, which helps them to formulate their next action. Dominguez et al. (1994) define SA as:

> *"Situation awareness is the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events."*

However, it is important to note that these definitions are developed for aviation, with a special focus on aerial combat. Thus the "elements" (Endsley, 1988) and the "environmental information" (Dominguez et al., 1994) in the two definitions refer the relative location of enemy combatants to a fighter pilot, weather conditions, and to the operational status of an aircraft (i.e. they compare to the kinds of layman's context discussed previously). However, in later years the concept of SA has also been applied to other fields such as human-computer interaction (Matheus, Kokar, & Baclawski, 2003), human-robot interaction (Yanco & Drury, 2004), and health-care (Cooper et al., 2010).

There are several aspects in both definitions for SA that are maybe equally important to social interaction between people. Both definitions categorize three layers of awareness: perception, comprehension, and projection. For perception, Endsley (1988) relates events to time and space. Thus, the "elements" that people can perceive as relevant for their situation are taken to happen in close spatial and temporal proximity. Dominguez et al. (1994) also stress a temporal element by saying that SA is "...the continuous extraction...". Therefore, people evaluate their SA in real-time, which is also what happens in social interaction. "Comprehension of their meaning" and "the integration of this information" are the resources through which people act in a situation. That is, people's actions are influenced by how they assign salience to ongoing events. The "Projection of statuses" and "anticipating future events" are also important in social interactions (Dominguez et al., 1994).

SA, as defined by Endsley (1988), can be used to describe aspects of social interaction, and it is also compatible with Clark's model of common ground. The ten elements of personal and communal common ground are analogous to the "elements in the environment". Comprehension and projection are also implicitly represented in Clark's model and can be observed in the assumptions people make in interactions. The assumptions people make about what common ground they share work as direct windows into how people evaluate the contextual elements of an interaction. In the following, I describe in more detail the relevance of SA for HRI research.

Situation Awareness in HRI

Much work on SA in HRI deals with a controller's SA when teleoperating robots. The focus here is on giving the controller a better 'picture' of where the robot is located in relation to points of interests or potential threats (Yanco & Drury, 2004). In these situations, the robot acts as a medium (Groom et al., 2011a) through which a controller can interact with a remote environment. This is useful in several contexts. For example, to gain access to areas that are simply dangerous to humans (Nonami, Shimoi, Huang, Komizo, & Uchida, 2000), in search-and-rescue operations (Dole, Sirkin, Currano, Murphy, & Nass, 2013), in communication over long distances (Adalgeirsson & Breazeal, 2010; Tanaka, Takahashi, Matsuzoe, Tazawa, & Morita, 2014), or to assist humans in complex operations such as surgery (Moustris, Mantelos, & Tzafestas, 2013). For example, Drury, Keyes, and Yanco (2007) compare situations in which a controller has access only to a digital map that is updated in real-time with situations in which a controller has access to a live video feed. They find that each of the methods gives access to different aspects of the situation. Other work also evaluates SA on the basis of control modalities (Adamides et al., 2017; Cross et al., 2009; Gómez, 2010; Kružić, Musić, & Stančić, 2017). Some researchers look into how operators' SA can be increased when controlling multiple robots (Crandall & Cummings, 2007; Cross et al., 2009; Envarli & Adams, 2005). Similarly, other researchers study how controllers can gain and maintain SA when teleoperating a robot (Hedayati, Walker, & Szafir, 2018; Johnson, Rae, Mutlu, & Takayama, 2015; Scholtz, Antonishek, & Young, 2005; Zheng, Glas, Kanda, Ishiguro, & Hagita, 2011).

One of the challenges that hamper discussions of SA to move beyond teleoperation is that most robots that can perceive many of the same elements that people can, but do not (yet) possess the ability to interpret the perception of these elements in the same way people do (Adams, 2005). Thus, most robots do not advance beyond Endsley's (1988) level 1 SA, while humans have access to all three levels (perception, comprehension, projection). For example, in one study, teams of robots could see each other, but the robots merely interpreted the signals as additional obstacles they had to avoid. Thus, to the robots, other robots were comparable with walls and debris (Drury, Scholtz, Yanco, et al., 2003).

Beyond Teleoperation

There is also some work that goes beyond teleoperation contexts by studying SA, for instance, in human-robot colloborative work. However, also within this context, much work focuses on the awareness of a human collaborator. For example, Scholtz (2003) defines four roles (supervisor, operator, mechanic and peer), which all rely on different aspects of SA. Thus, the robot is not seen as an agent that requires SA. This may be due to the fact that the cited work is from when HRI was still in its infancy, but even more recent work follows the same line of argumentation. For example, Dini et al. (2017) study how a human collaborator can anticipate when to turn his or her attention to the robot to complete a handover while being engaged in another task. Thus, the focus here is on how the human agent gains and maintains SA. Similarly, Unhelkar, Siu, and Shah (2014) compare human-human and human-robot teams to investigate performance and safety measures. While they do deal with SA, they do so strictly from the human team members' point of view. That is, the study addresses to what extent participants felt that they had an adequate awareness of what the robot was doing to feel safe. Furthermore, Côté, Canu, Bouzid, and Mouaddib (2012) present a robot controller that has the capacity to ask a human agent for help when it encounters a problem it cannot solve. They add features to the contoller's communication of the problem so that the human agent's SA is increased.

Situation Awareness in Robots

While much work deals with increasing humans' SA, there are also a few studies that implement SA in robotic systems. While they do not call it situation awareness, Pandey, Ali, and Alami (2013) implement *multi-state perspective-taking* into two different robotic systems in order to endow robots with the capacity to produce proactive behaviors based on what a human collaboration partner is currently doing and what he or she might do next. Proactivity relies on what an actor is likely to do next and can therefore be taken as a signal for SA. The robots in the study (Pandey et al., 2013) produce proactive behaviors during handover sequences. Results show that participants were less confused about what to do next, and they rated the robot as more aware and supportive when it produced proactive behaviors. In another study, a robot employed facial and skeletal tracking of the people it met in order to increase the SA of the robot (Mykoniatis, Angelopoulou, Schaefer, & Hancock, 2013). While the algorithms performed well, it is unclear from the study how the robot utilizes this information and what impact its SA has on interactions.

Even though technologies enable robots (and other technologies) to have access to an increasing number of sensors, there is only a relatively small body of research on how a robot can display its awareness to its human communication partners, and how these displays affect communication partners' perceptions and behaviors.

1.2.4 Indicators for Situation Awareness

In this section, I describe some possible ways in which, participants in interaction can indicate to each other what aspects of the situation they take into account.

Contingency

Contingency (Schegloff, 1968) is described as the property that binds conversational elements together so that they form a sequence. Sequence is to be understood here as a distinct sequential organization rather than as a set of events that are bound by temporality alone (1968, 1972). This is exemplified for example in adjacency pairs. In a question-answer exchange, the the question and the answer form a sequence, not because one comes after the other, but because the answer attends to one of more aspects of the question. Even when a question is not answered its absence becomes noticeable, which serves to show that an element of the sequence is indeed missing (Schegloff, 1968). The two elements of the sequence may be temporally disconnected, for example by insertion sequences, but they remain contingent (Schegloff, 1972). Contingency is, in CA, one of the basic premises for accomplishing interaction. It is also the reason why utterances or actions are rarely analyzed in isolation.

> "*Contingency - interactional contingency - is not a blemish on the smooth surface of discourse, or of talk-in-interaction more generally. It is endemic to it. It is its glory. It is what allows talk-in-interaction the flexibility and the robustness to serve as the enabling mechanism for the institutions of social life.*" (Schegloff, 1996)

As Schegloff (1972) notes, contingency extends well beyond the temporal and spatial proximity, through which adjacency pairs prototypically are exemplified. Rather, responses (or 'seconds' as they are known by in conversation analytical terminology) are given with a sensibility to the context (in a methodological sense of the word), rather than to whatever has merely temporarily preceded it. As Schegloff (1972) points out:

> *"These notes may be read as pertinent to some ways in which "contextual variation" affects interaction. It is being proposed that the much invoked "dependence, on context" must be investigated by showing that, and how, participants analyze context and use the product of their analysis in producing their interaction."*

Thus, contingency can be considered multivariate in the way that multiple contextual features can influence multiple interaction outcomes, but also in the way that contextual features are not only limited to talk, but can come in the form of multiple modalities (Hazel, Mortensen, & Rasmussen, 2014; Lindwall & Ekström, 2012; Mondada, 2009a; 2009b). Any one action never stands alone, but relies on context in both production and evaluation. This embeddedness between context (again, in an ethnomethodological sense) and individual actions is contingency. Interaction is thus "characterized by contingency at virtually every point" (Schegloff, 1996).

This conceptualization of contingency in CA differs somewhat from what is considered 'contingency' in HRI. Here, contingency is understood as a linear temporal relationship, in which behavior is influenced by a stimulus (Chu, Bullard, & Thomaz, 2014; Gold & Scassellati, 2006; Lohan et al., 2011). In HRI, contingency is understood in terms of cause and effect. This understanding stems from research in feedback models for infants (Gergely & Watson, 1999) in which researchers investigate the temporal relationship between stimulus and response in human infants. Therefore, the problem of contingency in HRI is often one of detection. The problem is solved by implementing sensors in robots that endow them with the ability to detect changes in human behavior to which they can produce a response. Thus, contingency in HRI is a feature that can be implemented, while in CA it is an ever-present feature of interaction. This is also part of the reason why analyses in CA are always presented on a case-by-case basis; the contingencies at play in any given interaction, and the understandings of context displayed, are individual. Only through careful detailed analysis can analysts uncover exactly what aspects of context communication partners attend to.

Both models of contingency rely on indicators for situation awareness. In both models, communication partners display an awareness to each others' action in their own conduct. The model of contingency, as understood in CA, explicates, on a very detailed level, the intricacies of social interaction, by making clear what aspects (context) of the interaction people attend to and how it affects their own conduct. As such, this model of contingency shows how interaction is accomplished and what resources people put to use in this accomplishment. The model used in HRI does not exhibit the same complexity, but therein perhaps lies its strength. Contingency, according to this model, is relatively easy to implement and operationalize, and the model also makes clear what features of conduct communication partners display an awareness towards.

Incrementality

Verbal communication is, as Schlangen and Skantze (2011) point out, almost always incremental. Participants in conversation produce speech in real-time, often without having a complete plan of what they are going to say or do during their turn (Brennan, 2000; Levinson, 2016; Skantze & Hjalmarsson, 2010). In interaction, people do not have fully formalized plans of actions before they carry out those actions. Rather, people constantly adapt to their surroundings and produce actions on-the-fly or change provisional plans as they see the need. For example, Suchman (1987) showed that people produce interactional contributions in a piecemeal fashion. That is, they produce their contributions (speech, gesture, etc.) in small chunks, constantly updating their contributions based on what is needed to accomplish the interaction. Evidence of this behavior is found, for example, in word searches and floor-keeping devices. These features would simply not be so prevalent in social interaction if people did not produce and process action incrementally. From a CA perspective, incrementality is uncontroversial (Goodwin, 1979). However, only over the last decade is incrementality being discussed in the field of HRI. Thus, to date many robotic systems do not process speech incrementally, although considerable work is being done to change this (see, for example Baumann, Kennington, Hough, and Schlangen (2017), Schlangen and Skantze (2011), Skantze and Hjalmarsson (2010)).

Incremental processing relates both to the production and comprehension of actions in social interaction (Schlangen & Skantze, 2011). Since incrementality can be seen as evidence that communication partners design their action based on ongoing changes in the context, incrementality works as an indicator for situation awareness, in which communication partners signal to each other that they attend to each others' behavior.

Proactivity

Proactivity relies on information what a communication partner might do next and can therefore be taken as a signal for SA. Proactivity is directed related to projectablity, which is part of Endsley's situation awareness model (Endsley, 1988). Projectability is the third and last level of the model and uses information from the first two levels, perception and comprehension, in order to forecast what will happen next. Therefore, displaying to communication partners an awareness of what is about to happen, displays an awareness of what has happened already, what action is currently ongoing, etc. Projectability is also a resource people use in interaction as it unfolds:

> *Sentential constructions are capable of being analysed in the course of their production by a party/hearer able to use such analyses to project their possible directions and completion and loci. In the course of its construction, any sentential unit will rapidly (in conversation) reveal projectable directions and conclusions[...].* (Sacks et al., 1974)

However, in CA, projectability (Sacks et al., 1974) and predictability (Liddicoat, 2004) primarily concern the turn-taking mechanism. In other words, people design their utterances to make the turn project when a turn-transition is coming up. These signals can be produced lexically, prosodically, or by using gaze, gestures or any other kind of modality available to communication partners.

Projectability is also implicitly part of Clark's (1996) model of common ground, in that when people signal to each other what their common ground is, they also signal what to expect. However, Clark never discusses projectability or proactivity specifically, other than saying that some events are *anticipated products*, based on people's intention (Clark, 1996, p. 22).

In HRI, proactivity is related to intention and intention recognition. The general idea is that a robot should attempt to find out what its human communication partner is doing, and on this background produce behaviors that support the human in his or her task (Ali, Alili, Warnier, & Alami, 2009). Proactive or anticipatory behaviors are therefore indicators for a robot's situation awareness and are direct clues through which communication partners can infer which aspects of the interaction robots take into consideration.

1.3 Aim and Structure of the Thesis

In HRI, context is dealt with only rarely. Context is, if invoked at all, often treated as a macro-level construct. For example, some research on context in HRI focuses on the 'cultural context' (Bartneck, Suzuki, Kanda, & Nomura, 2007; Trovato, Ham, Hashimoto, Ishii, & Takanishi, 2015; Wang, Rau, Evers, Robinson, & Hinds, 2010). Other works treat context on the interactional level, but reduce the concept to a certain type of activity (Read & Belpaeme, 2014; Salem, Ziadee, & Sakr, 2013) or to the goal of an activity (Nehaniv et al., 2005), while some also describe context as noise or as 'silent factors' (Cameron et al., 2015). Since context has no prominent role in the HRI literature, there is also only very little work that describes how situation awareness or robot's displays of awareness of contextual information affects human-robot interactions.

The aim of this thesis is to investigate how robots' displays of awareness of contextual features affect communication partners' perception and behavior. This investigation draws on common ground (Clark, 1996), situation awareness (Endsley, 1988) and on an ethnomethodological understanding of context. In the following, I describe how each of these resources are put to use in the thesis and how they relate to each other.

1.3.1 What does Context mean for this Thesis

Earlier in this chapter, I described three perspectives of context: layman's, ethnographic, and ethnomethodological. The understanding of context that is under investigation for this thesis is primarily what can be considered to be ethnomethodological context. This means that I investigate features of social interaction that signal an awareness of what is going on in the interaction and communicate what communication partners can consider common ground. While aspects of context, in an ethnomethodological sense, are what is considered for investigation, I also use an layman's and ethnographic understandings of context at specific places. Specifically, the effect of participant gender and the extent to which participants have interacted with robots before are considered for analysis in most of the statistical analyses presented in the thesis. The reason for this, rather un-ethnomethodological, choice rests in the fact that numerous studies in the HRI literature have found gender-related effects (e.g. Kuo et al. (2009), Lubold, Walker, and Pon-Barry (2016), Salem et al. (2013), Schermerhorn, Scheutz, and Crowell (2008), Siegel, Breazeal, and Norton (2009)) and some effects of participants' previous experience with robots (Dautenhahn et al., 2005; Koay et al., 2007; Salem et al., 2013; zu Borgsen, Bernotat, & Wachsmuth, 2017). However, these variables are not included in analyses to identify differences between, for example, men and women or experienced and naïve users. Rather, they are included in the analyses in order to control for effects that previous works indicate might be present. In other words, they are included in the statistical analyses and presentations to give a more precise picture of the effects of the experimental conditions.

1.3.2 Operationalization of Common Ground

The kind of common ground under investigation is generally personal common ground. While communal common ground is equally important to social interaction, and to the understanding of context, as personal common ground, robots are still so novel to the most of us that there is no communal common ground to establish or maintain. In other words, communal common ground builds on the assumption that there is a community, but as of yet no 'community' of robots exist. Thus I confine the analysis to aspects that are directly relevant for interaction and may be taken for granted in interaction between people. Within personal common ground, signals that communicate the perceptual and the actional basis, and personal diaries are investigated. The remaining two elements, 'acquaintedness' and 'personal lexicons' are not investigated as they would require an experimental design that spans several interactions over a relatively long period of time. Although such an investigation is relevant and will at some point become necessary, it is beyond the scope of this thesis.

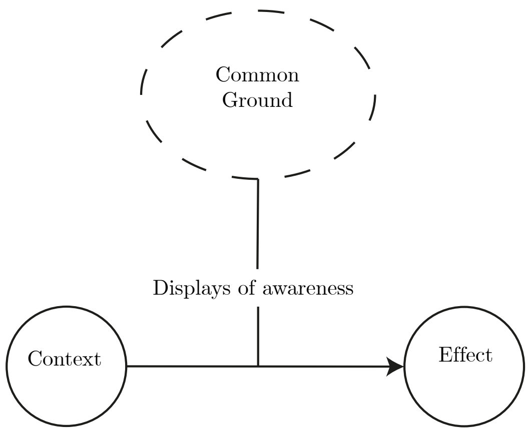


Figure 1.2: Conceptual Model

Context, common ground, and displays of awareness are combined in one conceptual model (visualized in Figure 1.2) that drives and informs the work presented throughout the thesis. Specifically, displays of awareness of contextual features communicate to partners what the (perceived) common ground is. In turn, the shared common ground between communication partners affects their perception of each other and how they interact. Common ground cannot be measured or analyzed directly, but what can be analyzed are the displays of awareness of context and how these displays are responded to (i.e. the effects). In the following, I describe how the three aspects of common ground are investigated through displays of awareness to contextual features.

Perceptual Basis

The perceptual basis comprises those elements in interaction that people can perceive and share (Clark, 1996, p. 112). This includes information such as the layout of the physical interaction space, the people in it, and events carried out in that space. The perceptual basis is thus a shared context to which all interaction partners can attend to. This is a huge problem for interactions with robots. People generally do not know how, what, or even if robots perceive anything, because robots do not possess a human nature (Clark, 1996, p. 106) that can guide what assumptions people can make of them. In ethnomethodological terms this means that robots are unlikely to be 'members' humans have encountered and dealt with before, and as a direct result thereof people cannot make use of their 'members' knowledge' to understand robots. The perceptual basis is usually not dealt with directly, but it is indicated in actions, for example through talk, gaze or gesture. Therefore, it is not these actions that themselves are part of the context, but they are indicators of a shared perceptual context. The perceptual basis is investigated in Chapter 4 and in Chapter 5, in which a robot produces verbal references to objects in the interaction space and comments on the state these objects are in. Through these actions, the robot signals to communication partners that these objects, and their states, can be considered common ground.

Actional Basis

The actional basis comprises joint actions people engage in (Clark, 1996, p. 114). In contrast to the perceptual basis, actional basis can be accessed directly, for example through talk. Talk can display to communication partners what aspects of the common ground they take for granted. The actional basis is usually expressed through speech, but can just as well be expressed through gestures or gaze behaviors. The actional basis is investigated through contingency, incrementality, and proactivity. Contingency is understood as a linear temporal relationship between cause and effect; i.e. using the understanding of contingency found in the HRI literature (e.g. Lohan et al. (2011)). Contingency works on multiple modalities that include, but are not limited to speech, gesture and gaze. The effects of contingency as an indicator for situation awareness are investigated using these three modalities. Chapter 3 investigates effects of contingent gazes and nods, Chapter 5 investigates the effects of contingent verbal responses and contingent face-tracking, and Chapter 8 investigates the effects of contingent responses to repair initiations.

Incrementality is generally seen as a model for speech processing (see e.g. Schlangen and Skantze (2011)), and this is also how I consider incrementality. Incremental speech is therefore speech that is produced in small 'chunks' or increments, where each new contribution is updated with information about the situation the speaker is in. Therefore, incrementality is a display of what contextual information a speaker considers to be relevant and thus signals to the communication partner what the speaker considers to be the common ground. Chapter 5 investigates verbal incremental feedback in a physical exercise, and Chapter 6 addresses the effects of incremental feedback in an object search task. In both of these studies, the verbal incremental feedback is designed to signal the robot's awareness of participants' nonverbal conduct.

Proactivity is an indicator for situation awareness in the sense that it signals to communication partners what is about to happen. Proactivity, as conceptualized and implemented for this thesis, does not include any measures of anticipation or intention recognition of the partner in interaction. Proactivity displays actors' awareness of their own actions as relevant context in communicating what they consider to be common ground. Thus, proactivity is an indirect signal that interaction partners orient to a joint plan. Proactivity is investigated in Chapter 7, in which a robot in a handover scenario proactively signals what it is going to do next using gaze cues.

Personal Diaries

Personal diaries refer to what has gone on previously in the current and in previous interactions already and which is part of the common ground that communication partners establish and maintain throughout the interaction (Clark, 1996, p. 114). Displays of memory show that actors not merely react to stimuli as they are perceived, but that actors can store the information and contextualize it, which is problematic for robots. For example, Christian (2011) posits that one way to distinguish between humans and machines is that humans can access and make use of past experiences and recontextualize them to new situations, while computer (and thus robots) cannot. Displays of memory therefore work as indicators to communication partners that previous interactions, or previous actions within the current interaction, can be taken as common ground. This use of memory is often taken for granted in interactions between people, but cannot be taken for granted in interactions between people and robots (see Christian (2011)). This aspect of common ground in HRI is investigated in Chapter 4 in which a robot displays an orientation to previous utterances and uses this information in its own productions.

1.3.3 Thesis Structure

The remainder of the thesis is structured as follows: In the following chapter, Methods & Data, I account for the methods I use and what possible consequences the choices of these methods may have for my results. The chapter also includes an overview of the empirical studies carried out. Following that chapter, I explore my research questions in a series of empirical investigations in Chapters 3 to 8. These empirical chapters are followed by a discussion of the results obtained and the insights they provide in relation to my research questions, to previous work, and with respect to what design implications can be drawn from the results. This discussion is followed by a conclusion, in which I attempt to combine all elements of the thesis and set a course for future work.

2. Methods & Data

In this chapter I present and review the methods used in data collection and analysis for the thesis. At the end of the chapter I also provide an overview of the empirical work that lay the foundation for the following six chapters.

2.1 Analytical Considerations

Historically, the field of Human-Robot Interaction was influenced to a large degree by computer science, the engineering sciences, and psychology. Especially psychology has influenced HRI methodically with its reliance on the 'hypothetico-deductive model' or the 'scientific method'. Thus, new knowledge in the field is generally derived from stating operationable hypotheses than can be tested in controlled experiments. Results are expressed in quantitative terms using inferential statistics, which then lay the foundation for the next set of hypotheses. I follow this tradition, but supplement quantitative analyses with qualitative analyses, in particular using ethnomethodological conversation analysis (Sacks et al., 1974).

2.1.1 Statistical Methods

Quantitative relationships are primarily evaluated using regression methods. Regression analysis has the advantage that it is very flexible in the type of data it can handle, and in the type of analyses it can perform. The type of regressions used throughout this thesis are multiple linear regression, which is used to analyze data with a numerical outcome (e.g. questionnaire responses), and logistic regression, which is used to analyze data with a binary outcome (e.g. the likelihood of one action over another). Results are, throughout the thesis, presented by visualization of means in bar charts. However, as most regression analyses are modeled with at least three predictor variables, the experimental condition, the participant gender, and the participants' previous experience with robots (on a four-point scale), the mean difference between experimental condition do not necessarily capture the complexity of the data. In order to better account for how multiple predictor variables affect the outcome regression are plotted in graphs similar to the one found in Figure 2.1.

Figure 2.1 shows how three three predictor variables (experimental condition, participant gender, and experience with robots) affect four outcome variables. Statistical significant relationships are denoted by fully marked lines (—), regardless of color, and marginally significant relationships are denoted by dashed lines (- - -). No lines means that no statistical relationship between variables can be observed. For each factor, one level is

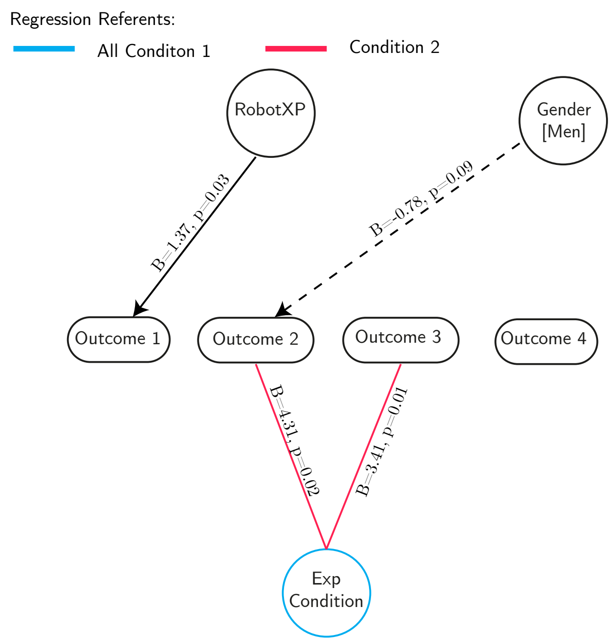


Figure 2.1: Modeled Regression

designated as the referent. In other words, the level that all other levels in the same factor is compared against. The referent is denoted by the lines between variables. For example, for the experimental condition, the referent is 'Condition 2', and for gender, it is 'women'. Therefore, for outcome 2 the regression coeffecient (B) for condition 1 is 4.31 in comparison to condition 2. Likewise, in outcome 2 the regression coeffecient for men is 0.78 in comparison to women.

2.1.2 Conversation Analysis

Conversational analysis (CA) is the study of social order in interactions between people. In CA, interaction is seen as jointly organized activity that is accomplished rather than produced. CA excels at uncovering the structure in interaction, by looking at the sequential unfolding of events and recurring patterns. One aspect that separates CA from other research methods is that in CA, analysts attempt to understand interactions as participants themselves understand them. In other words, interpretations of events are analyzed in the light of what understandings participants themselves make observable through their talk or conduct. As a research method, CA enables analysts to understand how people accomplish interaction as the interaction unfolds. This is a quality that is equally relevant in interactions with robots. Generally CA is used in HRI in two different ways; as a design resource, or as an analytical tool.

CA as a Design Resource

In one way, researchers review CA literature to discover how people engage in social action, and from this extract behaviors that can be implemented in robot designs. This is also very similar to the way psychology has informed HRI and other HCI-related fields. For example, in Pitsch et al. (2009), the authors implement behaviors on their robot based on two sociological concepts; ‘focused encounters’ (Goffman, 1961), and ‘the first five seconds’ (Schegloff, 1967). In an attempt to manipulate peoples’ gaze Kuzuoka et al. (2008) implement ‘restart’ and ‘pause’ behaviors on their robot, with a basis in insights reported by Goodwin (1980). The organization of turn-taking is a central element in CA methodology, and a number of studies have implemented features that relate to turn-taking in one way or another. For example, Yamazaki et al. (2008) implement a nodding feature timed to transition relevant places (TRP), which is a place in the interaction where a speaker change can occur (Sacks et al., 1974). Fischer, Lohan, Saunders, et al. (2013) implement certain contingent features into robots with a reference to how contingency contributes to joint-action (Schegloff, 1996). Aarestrup, Jensen, and Fischer (2015) draw inspiration from Pillet-Shore (2012) as they test how people respond to lexically and prosodically different robot greetings. In other studies, Oto, Feng, and Imai (2017) investigate how people deal with silence, with a special reference to pauses, gaps, and lapses, in interaction with a robot, and Ohshima, Fujimori, Tokunaga, Kaneko, and Mukawa (2017) have developed a conversational robot that designates next speakers in interaction. The authors’ modeled the turn-taking behavior of the robot after the ‘turn-taking’ system documented by Sacks et al. (1974). In an interaction scenario with a virtual agent, Muhl, Nagai, and Sagerer (2007) look at the role of trouble sources in interaction and investigate how people deal with it in interactions with their virtual robot. Other work use the notion of ‘recipient design’ (Sacks et al., 1974) when designing robot behaviors (Fischer, 2016b; Fischer, Lohan, & Foth, 2012), or use CA to model turn taking behavior in a robot system (Fukuda et al., 2016; Linssen et al., 2017; Okuno, Kanda, Imai, Ishiguro, & Hagita, 2009; Rossi, Ferland, & Tapus, 2017).

CA as an Analytical Tool

A second way in which CA is used is as an analytical tool for making sense of interactions between people and robots. Using CA in this manner is however not unproblematic. Conversation analysis is concerned with how people accomplish social action. The focus here is on the *how* and the *why* (Schegloff & Sacks, 1973). That is, in CA we look closely into which methods people use to make sense of the situation they are currently in. This is commonly referred to as the ‘members’ methods’, or ‘ethnomethods’. CA excels at uncovering what participants themselves find relevant and important in any interaction. That is, the analyst comes to see the social situation from the perspective of the participants, rather than from an *a priori* understanding of how the interaction should proceed, by looking only at the observable actions performed, and how they are responded to. Membership categories, such as gender, occupation, and age also only become relevant if participants orient to them through their social conduct. Some conversation analysts might argue that using CA to investigate interactions between people and robots makes only little sense, as robots are not ‘members’ in a sociological understanding[1]. However, here it is important to note that as analysts we are not interested in robots’ ‘methods’ - these have already been designed either explicitly through a script or one or more algorithms. Instead, what we are interested in is how people respond in-the-moment to the interactions that we

[1]Recent work in CA, for example, also study human-animal interactions (Mondémé, 2011).

design. One of the first to look at human-machine interactions from an ethnomethodological perspective was Lucy Suchman (1987), who investigated interactions between people and a photocopying machine. What was innovative about her approach was that she looked at not only at which resources people had access to and which understandings they made relevant, but also showed which resources the machine had access to, and how peoples' actions were interpreted by the machine. In other words, she conceptualized the machine as a participant in interaction.

CA, unlike other data-driven methods, is not driven by hypotheses or formal models of how interaction work or should work. Instead, insights are drawn and developed from the data (usually video recordings) at hand, and by looking very closely at what participants' themselves make relevant. One of the key methodical concepts behind the approach is 'unmotivated looking' (Sacks, 1984), which means to look for new phenomena, free of preconceptions and hypotheses. Once one or several candidate phenomena have been discovered, the 'looking' becomes motivated as researchers try to find out whether the same phenomenon or social practice can be found under different conditions (i.e. in several cases, with different people). As the work proceeds, findings are gathered and sorted into collections, which also serves as an indicator for how common the phenomenon under investigation is. However, results are never validated statistically in the same way that hypotheses are (in)validated using inferential statistics. Thus, CA studies published in HRI are often labeled 'case studies' (e.g. Arend and Sunnen, 2017; Dickerson, Robins, and Dautenhahn, 2013; Pitsch and Koch, 2010; Robins, Dautenhahn, and Dickerson, 2009).

One way to get around this problem is to publish quantitative results alongside the qualitative results. This can be done in (at least) two ways. In one way, the 'unmotivated' looking and the subsequent analysis identifies phenomena that can be coded, and later processed statistically. This is what Gehle, Pitsch, Dankert, and Wrede (2017) refer to as 'Conversation Analysis with quantification'. For example, they look at when people are engaged in interaction with a robot in a museum. Based on this qualitative analysis, they classified under which circumstances people engaged with the robot, taking gaze, body posture and distance to robot into account. Similarly, Pitsch et al. (2009) identify cases in which a robot is able to engage people in what they call a "contingent entry", which they classify as entry into an interaction where the robot responds appropriately and timely to the user, and the user responds appropriately and timely to the robot. They then code the number of people who respond to the robot when it is nodding, speaking, whistling, shifting positions, or whether they leave the interaction prematurely. Finally, they compare the people who entered the interaction contingently with those who did not, using inferential statistics. Similarly, Gehle et al. (2015) qualitatively identify trouble sources in interactions between visitors to a museum and a museum guide robot, which then are quantitatively processed. Similar approaches to combining CA and inferential statistics are also utilized by Opfermann and Pitsch (2017) and (Cyra & Pitsch, 2017). Finally, also Fischer et al. (2015) start with 'unmotivated looking' in a collaborative assembly scenario. Here they identified participant responses, such as nodding, smiling, and gaze behaviors using conversation analytical methods. Subsequently, these responses were coded and processed with inferential statistics.

Another way in which CA and quantitative methods are joined is in hypothesis-driven analyses, in which CA-analyses are used to explain and interpret results of quantitative analyses. For example, Pitsch and Wrede (2014) carry out an experiment with a robot in a museum, in which the robot attempts to attract and keep visitors' focus and attention to the object that the robot refers to, using different coordination methods for talk and gestures. Initially, they find that success varies between 27 and 95%, but post-hoc analyses of the interactions reveals that visitors experienced problem when they were not facing the robot at the precise moment of a deictic gesture (Pitsch & Wrede, 2014). Also Wrede, Buschkaemper, Muhl, and Rohlfing (2006) use statistical methods together with conversation analysis. They designed an experiment with their robot BIRON, in which users interacted with either an introverted or extroverted version of the robot. The questionnaire analysis shows that participants prefer to interact with the extroverted robot, and the CA-analysis shows that the extroverted robot provides more feedback and thus "provides more access to the user" (Wrede et al., 2006).

There are however, also examples of work in which interactions between people and robots are analyzed using CA without quantification, and without being labeled as 'case studies'. Rather the respective researchers build 'collections'. That is, instead of subjecting a particular phenomenon to statistical processing, they attempt to find other instances of the same phenomenon in the data. During this process phenomenons are further refined and formalized (Hutchby & Wooffitt, 2008, pp. 89–90). Here, statistical processing becomes irrelevant as there is no predefined hypothesis and therefore nothing to compare quantitatively. An example of this kind of work is found in Pitsch, Vollmer, and Mühlig (2013), in which the authors trace how people respond to different kinds of robot feedback. Muhl and Nagai (2007) classifies in an ethnomethodological perspective, the different responses people display when they experience trouble in interaction. Likewise, Plurkowski, Chu, and Vinkhuyzen (2011) investigate interactional in an HRI scenario, with only very little quantification.

2.2 Data collection methods

Data for all six empirical studies are carried out as controlled experiments with clear separation of experimental conditions.

2.2.1 Wizard of Oz

All studies presented in this thesis are carried with some aspect of human teleoperation, also known as Wizard of Oz (Weiss et al., 2009), which is a method that has been used extensively in Human-Computer Interaction research, usability engineering, experimental psychology and now increasingly in Human-Robot Interaction. The method refers to an experimental setting in which participants interact with a piece of technology (in this case a robot) believing it to be autonomous, while it is in fact operated by an experimenter who is hidden away from the participants. It is thus very similar to the Wizard of Oz from the story *The Wonderful Wizard of Oz*, where a man hides behind a black curtain and pretends to be a powerful wizard through the use of technology. Although the method has come under some critique (Riek, 2012) it remains a widely accepted method in both the human-computer interaction and human-robot interaction research communities.

In the current investigation, the human teleoperator will be referred to as the 'wizard'. Once the experiment has ended participants are debriefed and told how the technology works and also introduced to the person who was controlling the robot (the wizard), if they were interested.

2.3 Responsible Conduct of Research

Data for the project consist of three parts. One part is demographic information of participants' attitudes and their previous experience interacting with robots and other technologies. The only identifier is participants' ID which is a number assigned to them when they sign up to participate. The second part is participant responses to post-experiment questionnaires. The questionnaires are generally subjective ratings of a robot's performance or behavior. Again, no other identifying data are saved with these responses other than the participants' assigned ID number. Finally, the third part consists of video recordings from the experiments. Aside from these parts participants also fill out an informed consent before taking part in any experiment. The form includes their name, their assigned participant ID, and a checklist with which participants can choose how I may use the data and to whom I may show it. This is the only document which links the assigned participant ID together with the participant name. All the data are stored on a university server through which I control the access. All participants are given chocolate or cookies as compensation for their time.

2.4 Data

Data for the thesis are collected in several controlled experiments using three different different robot platforms. A short description of each experiment is provided in the following.

2.4.1 Study 1: Contingent Gaze

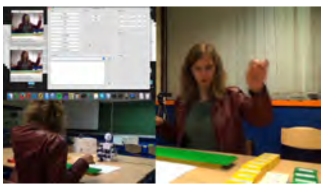


Figure 2.2:
Study 1: Contingent Gaze

In this experiment 28, students and staff from the Vrieje Universiteit Brussels interacted with the EZ-Robot for about 10 minutes each. At their disposal they had a collection of Lego Duplo blocks with word printed on them, which the robot was be able to read out loud. This study investigates the contextual features *contingent gaze*, and *head behavior*. The experiment has three experimental conditions: *contingent gaze*, *contingent nods* and *random gaze* behavior. Analytically, the study evaluates participants perceptions', language use, and gaze behavior.

2.4.2 Study 2: The Discourse Record

In this experiment, 52 students from the University of Southern Denmark interacted with the EZ-Robot for about 10 minutes. The robot instructed participants to build a frog out of Lego blocks. The experiment manipulated situation awareness in two conditions. In one condition, the robot commented on the weather and made reference to utterances made earlier by the participants. Thus, the study investigates the contextual features *dialog history*, and *the perceptual basis*. Only perceptive data are collected for this experiment, which is evaluated through an analysis of post experiment questionnaires.

Figure 2.3: Study 2: The Discourse Record

2.4.3 Study 3: The Perceptual Basis, Face-tracking, & Incrementality

Figure 2.4: Study 3: The Perceptual Basis, Face-tracking, & Incrementality

In this experiment, 80 students and staff from the University of Southern Denmark interacted with the EZ-Robot for about 10 minutes. Participants were asked to do a series of exercises by the robot. The experiment features four different experimental conditions, enabling analyses of different aspects of situations awareness, non-verbal contingency and incrementality. The contextual features under investigation in this study are *the perceptual basis*, *face-tracking*, *incrementality*, and *global contextual awareness*. Interactions are evaluated through analysis of post experiment questionnaires and objective behavioral measures implemented into the experiment.

2.4.4 Study 4: Incrementality

In this experiment, 51 students and staff from the University of Southern Denmark interacted with the Turtlebot for about 10 minutes. The robot led participants through a lab, asking them to pick up items along the way. The experiment manipulates situation awareness on the basis of incremental speech. On two occasions during the tour, participants need to pick up an item that is hidden from plain sight. Here, the robot employs word-level incrementality to guide participants in one direction or the other. The contextual feature under investigation is *incrementality*. The experiment is evaluated through an analysis of post experiment questionnaires, and through an objective measure of task performance.

Figure 2.5: Study 4: Incrementality

2.4.5 Study 5: Proactivity

Figure 2.6:
Study 5: Proactivity

In this experiment, 85 participants interacted with an industrial robot to collaborate on the construction of a piece of furniture. In the experiment, participants were told to assemble an IKEA children's stool with the assistance of the robot. Their task was to instruct the robot to fetch the legs of the stool, while the participants themselves had to perform the actual assembly. It was left open to the participants exactly how to instruct the robot. The instruction consisted of two phases: a fetching phase, in which participants had to indicate to the robot which of the four legs they wanted, and a handover phase in which the participant had to let the robot know where to to deliver the leg. The participants then connected the leg to the seat. The robot employed either reactive or proactive gaze between phases. The contextual feature under investigation for this study is *proactivity*. The study investigates participants' perception, through post experiment questionnaire, and gesture and gaze behaviors.

2.4.6 Study 6: Contingent Repair

This study used an identical setup to the study presented above. In this experiment, the robot had the ability to adjust its behavior if directed to do so by the participants. That is, when the robot failed to recognize the right stool leg to pick up, participants could direct it to the right one. In the second condition, the robot ignored participants' attempts at correcting its behavior. The contextual feature under investigation for this study is *contingent repair*. This is investigated by analyzing participants' perception of the robot, and by analyzing how participants perform handovers.

Figure 2.7:
Study 6: Contingent Repair

3. Study 1: Contingent Gaze

3.1 Introduction

Gaze[1] is an integral part of social interaction like speech, gesturing, and other semiotic resources. In fact, Argyle and Cook (1976, p. 167) posit that "any account of social behavior that fails to mention gaze is completely inadequate". Therefore it is only natural that robots designed for social interaction with humans should be able to recognize, process and produce human-like gaze behaviors. Researchers in HRI scour findings from studies of social interaction and human behavior in order to find features that can potentially advance robust and credible human-robot interactions. A wide range of different gaze behaviors have already been identified and implemented in robotic systems. This includes (but is not limited to) gaze aversion (Andrist, Tan, Gleicher, & Mutlu, 2014), pro-active/anticipatory gaze (Huang & Mutlu, 2016; Moon et al., 2014), mutual gaze (Ishii, Nakano, & Nishida, 2013; Richter et al., 2016), contingent gaze (Lohan et al., 2012; Skantze, Hjalmarsson, & Oertel, 2014; Yu, Scheutz, & Schermerhorn, 2010), gaze to manage turn-taking (Fischer et al., 2015; Mutlu, Shiwa, Kanda, Ishiguro, & Hagita, 2009) and gaze in groups (Pitsch & Gehle, 2013).

Contingent gaze, also known as deictic gaze (Argyle & Cook, 1976) or gaze following (Yu et al., 2010), is especially interesting as it offers an online view of what interaction partners find salient at any given point during the interaction. Contingent gaze is a direct indicator for a robot's situation awareness. Contingent gaze in HRI, and the way it is referred to in this chapter, means that gaze of one actor exists in a temporal cause-and-effect relationship with the gaze of another actor (Lohan et al., 2012). In HRI, Contingent gaze is usually validated by running controlled experiments in which naïve participants interact with a robot that uses contingent gaze. In these experiments, there is usually a control condition, in which participants interact with a robot that uses static or random gaze. However, contingent gaze is rarely ever compared to other forms of contingent responses, for example contingent nods. This makes it difficult to ascertain what exactly the contribution of contingent gaze is. Is contingent gaze a unique feature in HRI, or can other signals also perform the same social function? Does contingent gaze successfully perform or contribute to certain social actions because it is contingent, or is it successful because the combination of gaze and contingency creates a unique social signal?

This question is inspired by earlier work on meaning-making in sociology and HRI. Garfinkel

[1]Most of programming for this experiment was carried out by Katrin Lohan and Ingo Kellner from Herriot-Watt University. Data collection was done with the assistance of Hoang-Long Cao and Yong Ding at Vrieje Universiteit Brussels

(1967) showed in his "counselor experiment" that people rationalized and perceived responses as meaningful even when presented with contradictions by a 'counselor' with which they were interacting through a text interface. Garfinkel asked his students to think of a problem they had, describe the problem, and pose questions that could be answered with *yes* or *no* to a counselor who was situated in an adjoining room. What the students did not know was that the counselor randomly answered the students' question with *yes* or *no*. Garfinkel's experiment shows that people are quite flexible when it comes to making sense of what their communication partner says as something that is meaningful and coherent. Likewise, Fischer (2006) set up an experiment in which students were to schedule an appointment on a computer system using natural language. Unknown to the students was that the system responses were pre-synthesized and played according to a fixed script which was ignorant of their utterances. Yet, the students rationalized the system's behavior, perceived them to be meaningful and appropriate, and attempted to find meaning where there inherently was none. A reasonable explanation for why people find responses in the experiments meaningful is that both the 'counselor' and the 'appointment system' responded timely to their queries, and as a result participants assumed that they were meaningful.

The studies by Garfinkel and Fischer indicate that the timing component of contingent responses plays a prominent part in how people understand a response. Do timing play an equally prominent role when looking at contingent non-verbal responses? The current study explores this question in a controlled experiment in which participants interact with a robot that uses either contingent gaze, contingent nods or random gaze. The study deals specifically with two research questions. First, how do contingent non-verbal responses affect how participants respond, adjust to, and perceive the robot in comparison to random responses, and second, how do participants respond, adjust and perceive the robot differently between the two contingent conditions? The study relies on two kinds of data. First, post-experiment questionnaires were used to evaulate participants' perception of the robot. Second, interactions/experiments were video-recorded as to allow for analyses of the robot's and the participants' linguistic productions and gaze behavior.

3.2 Previous Work

Previous work concerns studies of the effects of contingent gaze in interactions between people and in interaction between people and robots.

3.2.1 Effects of Contingent Gaze in Interactions between People

We know from studies of gaze between people that gaze is variable (Argyle & Graham, 1976), participants in interaction respond contingently to each other's gaze signals. In his study on gaze direction, Kendon (1967) reports that gaze behavior varies from person to person, but that people coordinate their gaze behavior very tightly with their communication partners. That is, while much interpersonal variation can be observed in peoples' gaze behavior, people in interaction come to an 'agreement' as to how long they gaze at each during utterances, during listening, and during silences. This 'agreement' comes about in various ways and displays some of the functions contingent gaze fulfills in human communication.

Specifically, gaze is used as an interactional resource in turn-taking (Sacks et al., 1974), when initiating and completing repair (Goodwin, 1980), in order to provide and receive feedback (Argyle, Ingham, Alkema, & McCallin, 1973; Goodwin, 2000) and to modulate intimacy (Abele, 1986; Adams Jr & Kleck, 2005; Argyle & Dean, 1965).

This 'agreement', as Kendon (1967) describes it, has more recently been characterized as "a complex interactional dance, as it were, with frequently alternating periods of gazing at the other and gazing away" (Kendrick & Holler, 2017). In their study of gaze in interaction, they find that people tend to avert gaze when delivering a dispreferred second. Therefore, the 'conditionally relevant' next action also affects people's gaze behavior. These results are nicely in line with findings by Argyle and Cook (1976) who report that people who cooperate or people who like each other are more likely to establish mutual gaze. Given that dispreferred seconds are displays of disaffiliation and often accompanied with interactional trouble, it follows that people are less likely to establish and maintain mutual gaze.

These studies show some of the contingencies that are in play during interaction between people. Since studies have shown HRI to be similar to human-human interaction in certain ways, it is reasonable to assume that a robot that is able respond to some of these contingencies may be more likely to come across as more engaging and responsive as one that does not.

3.2.2 Effects of Contingent Gaze in HRI

In recent years, several studies have investigated the effects of contingent gaze in HRI. In HRI, Gaze has, as in human interaction, been found to facilitate turn-taking (Lallee et al., 2013; Mutlu, Kanda, Forlizzi, Hodgins, & Ishiguro, 2012; Richter et al., 2016). That is, in an HRI experiment in which the robot did not talk, Lallee et al. (2013) found that participants were still able to organize smooth turn-taking by using gaze. What is more, others have found gaze to be a reliable resource in facilitating turn-taking even when there is talk involved (Mutlu et al., 2012). Richter et al. (2016) finds that mutual gaze produced at the end of an utterance in multiparty conversations with a robot is a meaningful turn-yielding cue. Jokinen, Harada, Nishida, and Yamamoto (2010) trained an SVM classifer to predict when participants change turns in interaction. Feeding the classifier with utterances, speech act types, and gaze behavior, the classifier is reported to predict when a speaker change occurs in interaction. However, the authors have not tested the system in a live scenario with a robot. People have also been found to reciprocate robot's gaze behaviors (Xu, Zhang, & Yu, 2016).

Others have defined and implemented contingent gaze systems. For example, Lohan et al. (2011) developed a contingent gaze module (reported on in more detail in Lohan et al. (2012)) with which a robot can infer the gaze direction of communication partners. They tested this module in a user study and found that participants interacting with the robot equipped with the contingency module direct their attention more to the robot's conduct than participants interacting with a non-contingent robot, who focused their attention mainly on the objects they both were handling. Fischer, Lohan, Saunders, et al. (2013) expand on this work with linguistic utterances of users who interact with a robot that either uses its gaze to track objects in the work space or uses the contingency module described in Lohan et al. (2011). Results show that when the robot uses the contingent gaze module, participants structure their utterances to a greater extent, reduce the complexity of their utterances, and trust the robot more to learn from the interaction. That is, they take the discourse record into account. In another study using the same contingency module, Fischer, Lohan, Nehaniv, and Lehmann (2013) investigate how a robot's learning progress affects participants' behavior when engaging with the robot. They find that only when the robot utilizes the contingent gaze module do participants adjust their behavior to the robot. Similar results are reported by Pitsch et al. (2013). These studies show that participants' ways of engaging with a robot changes (i.e. participants employ different methods) as the robot displays situation awareness to different aspects of the participants' behavior.

Another approach to studying contingent gaze in HRI was developed by Pitsch et al. (2009). Their robot responds contingently to participants' gazes. Specifically, their robot was a museum guide, programmed to observe participants' gaze and respond to it. They found that when the robot is able to initially respond contingently to participants' gaze, participants were less likely to leave the interaction prematurely, and more likely to respond to the robot's degreeting at the end of the interaction. Another approach to study the effects of contingent gaze is presented by Carlmeyer, Schlangen, and Wrede (2016a) who work with a virtual agent. In the contingent condition, their virtual agent initiated a self-repair by self-interrupting in an attempt to regain the participant's gaze when the participants were inattentive. Results showed that the virtual agent equipped with contingent gaze was rated as significantly less nice, pleasant, likable, and friendly.

Mehlmann et al. (2014) report on a study in which their robot was equipped with what they call referential gaze. Referential gaze is here comparable to contingent gaze, but rather than relying on a participant's gaze for the robot's own gaze behavior, the robot tracks objects in the joint work space. They report that employing referential gaze, their robot was seen as more natural, pleasant and efficient in comparison to when it did not. However, along with the referential gaze they also test the effects of social gaze, which are gaze cues produced by the robot at transition relevant places. The robot followed participants' gaze and hand movement to signal attentiveness. In the non-social gaze condition, the robot generally looked at the workspace 70% of the time and at the user 30% of the time. It did not respond to turn-yielding gaze cues but took its turn always one second after the participant had finished his or her turn. Surprisingly, results show that the non-social gaze outperforms social gaze on all metrics. Especially the timing of the robot taking the turn was considered more 'appropriate'. These results indicate that the timing aspect of contingent action can in some cases take precedence over other components in contingent action. That is, doing 'something' at the right time is better than doing the 'right thing' at an 'inappropriate time'.

Only a few studies deal with how random gaze affects participants perception and engagements with robots. Random gaze refers here to gaze behaviors that are not systematized, and not contingent on any particular behavior. That is, there is no systematic order that governs the gaze behavior, and the behavior is not (designed to be) informative. One of the studies that do deal with random gaze (Skantze et al., 2014), compare contingent[2] with random gaze in a map drawing exercise. The study finds that participants find the contingent gaze more informative and felt better equipped to solve the task when interacting with the robot using contingent gaze. Similar to Fischer et al. (2012) and Fischer et al. (2012) Skantze et al. (2014) also report that participants produce fewer utterances in the contingent condition. Finally, the authors report that the robot in the random condition produced more restarts, which the authors take as evidence that participants had more trouble following the instructions. Another study also comparing contingent gaze with random gaze (Yu et al., 2010) found that participants produce more tokens and more words per utterance in the random gaze condition[3]. Thus, the finding that people reduce the linguistic complexity of their utterances when interacting with a robot using contingent gaze (also reported by Fischer, Lohan, Nehaniv, and Lehmann (2013), Fischer et al. (2012) and Skantze et al. (2014)) seems to hold across studies and robots.

How people's gaze behaviors are influenced when interacting with a robot using contingent gaze has also been investigated to some degree. However, likely due to the multifunctional nature of gaze, results are not coherent. For example some studies report that participants gaze more to a robot that uses contingent gaze in comparison to a robot that uses random gaze (Xu, Li, & Wang, 2013), while others report that participants gaze more towards the robot in a random gaze condition (Yu et al., 2010).

3.2.3 Effects of Contingent Nods in HRI

Nods between people are extremely multifunctional in interaction. They can function as turn-taking devices, display agreement or work as non-verbal backchannels for instance (Schegloff, 1987a). Therefore several studies investigate how nods can be produced on different robotic platforms. For example, Ishi, Liu, Ishiguro, and Hagita (2010) develop different methods of producing head nods and investigate how 'natural' they seem to people. Liu, Ishi, Ishiguro, and Hagita (2012) build upon this work and extends the model to further increase 'naturalness'. Likewise, Yamazaki et al. (2008) investigate when head nods should be produced in relation to speech. Also Schroder et al. (2012) emphasize the importance of head nods in interaction, in their system, the Sensitive Artificial Listener (SAL), as do Skantze, Johansson, and Beskow (2015) in the development of the Furhat robot. Sidner, Lee, Morency, and Forlines (2006) built a robot that is able to recognize head nods from humans. Common for all of the above studies is that they take for granted that head nods improve and are necessary for smooth human-robot interaction, without investigating the effects in detail.

Krogsager, Segato, and Rehm (2014) work from the assumption that the back-channeling function of nods in conversation incites communication partners to continue speaking, and thereby maintains their engagement. They report on findings from a controlled experiment in which participants interact with a robot that either nods while they are talking, or does not nod. However, they find that the non-nodding robot elicits more utterances

[2] The authors refer to the contingent condition as the consistent condition, but the condition is comparable with the general definition of contingent gaze

[3] The authors refer to the contingent gaze condition here as the 'following' condition

from the participants than the nodding robot. They follow up on this work by replicating the study, using a virtual agent (modeled after the same robot). Here, they find that participants speak much longer when the robot is nodding, than when it is not. The authors therefore conclude that verbosity is modified by embodiment. This conclusion is congruent with other work, which finds that people produce shorter (but more) utterances when interacting with a robot that can move both eyes and head compared to a virtual agent, which produces the same behavior (Fischer et al., 2012).

In a different study Riek, Paul, and Robinson (2010) investigate the effects of mimicry. They present a study in which a robot either mimics all head gestures produced by the participants in real-time, mimics only vertical head nods, or none at all. They find no significant differences between conditions on participants' subjective ratings of the satisfaction of the interaction. In an early study of responsiveness in virtual agents, Gratch et al. (2006) report on an experimental study in which in one condition a virtual agent responds to changes in speaker pitch, loudness and nods, by nodding. The robot was also able to nod in the control condition but did this (along with other behaviors) in a random fashion. The results show that participants in the responsive condition speak more and longer with the robot. Although the use of contingent nods is just one of several behaviors available to the virtual agent in the responsive condition, it indicates that contingent nods have a positive effect on engagement. In a different study Jung et al. (2013) find that engagement increases when robots use nodding as a back channeling device, but that they are rated as less competent.

However, it is as of yet, based on the literature review above, not possible to tell whether the positive perceptive effects are especifically related to nods, or related to the fact that 'something' happens (as suspected by (Mehlmann et al., 2014)).

3.2.4 Summary of the Literature Review

Contingent gaze and contingent nods share some functionalities, such as managing turn-taking in conversation. However, they also contribute each with different functionalities in conversation. Contingent gaze shows great promise as a method to maintain and even increase engagement, but it is as of yet still unclear how exactly contingent gaze contributes to interactional engagement and how much of the potential contribution is due to the 'right' timing of virtually any other action - such as nods. The current study investigates this interaction by comparing two contingent actions - gaze and nod - with a control condition in which the head movement and gaze of the robot is random.

The above review of the current literature allows for the formulation of a set of testable hypotheses. Considering that previous work has approached the analysis of the effect of contingency through analyses of questionnaire responses, linguistic production, and gaze behavior, the current study will likewise formulate and test hypotheses in these three domains. This ensures that results gained in the current study are comparable to previous research. Moreover, by collecting data about participants in these three ways allows for analyses that combine subjective and objective measures.

Contingency and timing have been found to lead to increased sociality and increased perceived competence. As different contingent actions are only seldomly compared directly, it is unclear how the two contingent conditions will differ. However, some works indicate that the right timing is more important than the content of the action (Fischer, 2006; Garfinkel, 1967; Mehlmann et al., 2014). Studies on peoples' linguistic production in interactions with a robot (however few) suggest that people reduce the linguistic complexity by producing more but shorter utterances and explain complex concepts by using expository utterances when interacting with a robot that employs contingent gaze (Fischer, Lohan, Saunders, et al., 2013). In addition, linguistic features that describe the interpersonal relationship between speakers, such as personal pronouns and tag questions have been found to occur more often in interactions with robots using contingent gaze (Fischer, Lohan, Saunders, et al., 2013).

Not much work has been done on people's gaze behavior when interacting with a robot using contingent gaze and/or nods. However, from studies of interaction between people (which must be considered to be contingent) we know that people often make gaze shifts and usually maintain mutual gaze about 30% of the time (Argyle & Graham, 1976). Considering that the two contingent conditions only differ in the modality of their contingency, it is still unclear how this will affect participants' gaze behavior. While there is ample evidence to support the argument that contingent gaze leads to many desired effects, such as increased sociality, recipient design, and more robust interactions, it is still unclear exactly how timing and contingency interact.

3.3 Method

A language game was designed, in which participants interact with a robot in a session in which they tutored the robot on the rules of word order in English. Specifically, participants are asked to teach a robot how sentences are structured in English. The experiment consists of two phases; an introductory phase in which participants explain in their own words what the basic linguistic constituents of a sentence are, and an instructional phase in which participants construct three sentences for the robot, each varying in complexity. At their disposal they have a collection of Lego Duplo blocks with words printed on them, with which they can demonstrate how sentences are formed (see Figure 3.1 and Figure 3.3).

3.4 Procedure

Participants were first given a text which explained the basics of English syntax (see Appendix C.1). The introductory text was used to ensure that all participants start with the same minimum base knowledge. Participants were also given a list of tasks to do (see Appendix C.2). While the tasks define the complexity of each sentence, it was up to the participants what word they put in each position. For example, the first task asked participants to construct a sentence using just a subject, a verb, and an object. The participants were then free to choose from the around 80 word blocks to form a sentence using those elements. The robot was pre-programmed to respond verbally and non-verbally to a number of non-verbal actions performed by the participant. It could also engage in small conversational exchanges, but only at the mercy of a 'Wizard of Oz' (Weiss et al., 2009), located in an adjacent room (see Section 3.4.2 below for an overview of the protocol).

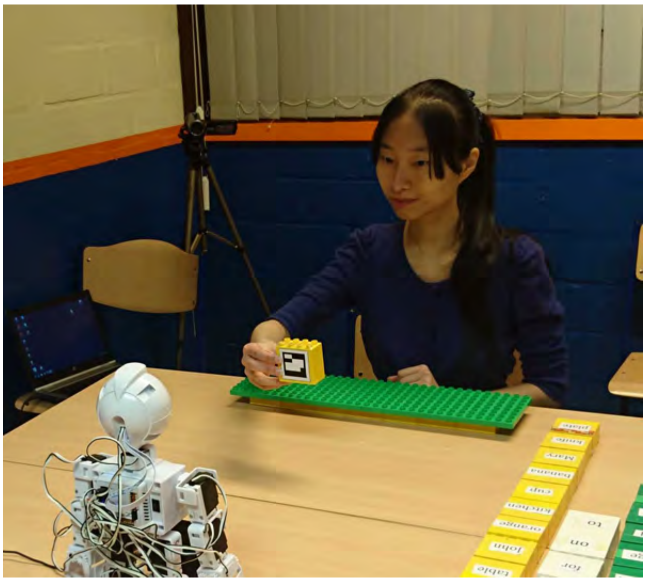


Figure 3.1: Participant shows one block to robot

3.4.1 Experimental Conditions

The study was set up in a between subject experimental design with three conditions. In the *contingent gaze* condition, the robot followed the gaze direction of the participant, so that when a participant gazed towards a certain block, so did the robot, just as it pointed toward an object if the participant picked one up. In the *contingent nod* condition the robot's gaze was static but responded to the participants' verbal and nonverbal actions by nodding[4]. In this condition, the robot's gaze remained fixed in one position. In the *random gaze* condition, the robot's gaze fluctuated in a random fashion. That is, the robot does not respond to the participant's gaze behavior.

3.4.2 Interaction Protocol

First, participants had to explain in their own words how to construct sentences in English. The robot was able to read out words or complete sentences when blocks were in its field of vision. Users were told that the robot understood basic English, but that it could learn more when instructed by a human tutor. Furthermore, the robot had the ability to read words out loud when these were shown to it. Users were instructed to explain the:

- syntactical functions of the different clause elements (verb phrase, subject, object, and predicate).
- basic word order of English in sentences containing a subject, a verb phrase and an object and by demonstrating using examples.
- basic word order of English in sentences containing a subject, a verb phrase, an object and a predicative.

[4] Nodding only occurred in the *contingent nod* condition

- basic word order of English in sentences containing a subject, a verb phrase, an object and an adverbial phrase.

The robot was controlled remotely, using the Wizard-of-Oz methodology (WoZ) from an adjacent office and followed a semi-set script. However, the controller was also able to respond to spontaneous prompts from the user. In addition, the robot could say ‘okay’ to signal understanding, and it was able to answer to yes and no questions (wizard controlled). There was no pre-scripted dialog or action protocol; instead the robot responded to the users’ conduct. The robot read a word out loud when a user was holding up a ‘word’ block or an entire sentence on the table when this was formed. One might fear when using WoZ that people are not really interacting with the robot, but rather use it as a proxy to interact with another human (Riek, 2012). However, during the debriefing all participants expressed that they were surprised that the robot was operated by a human controller.

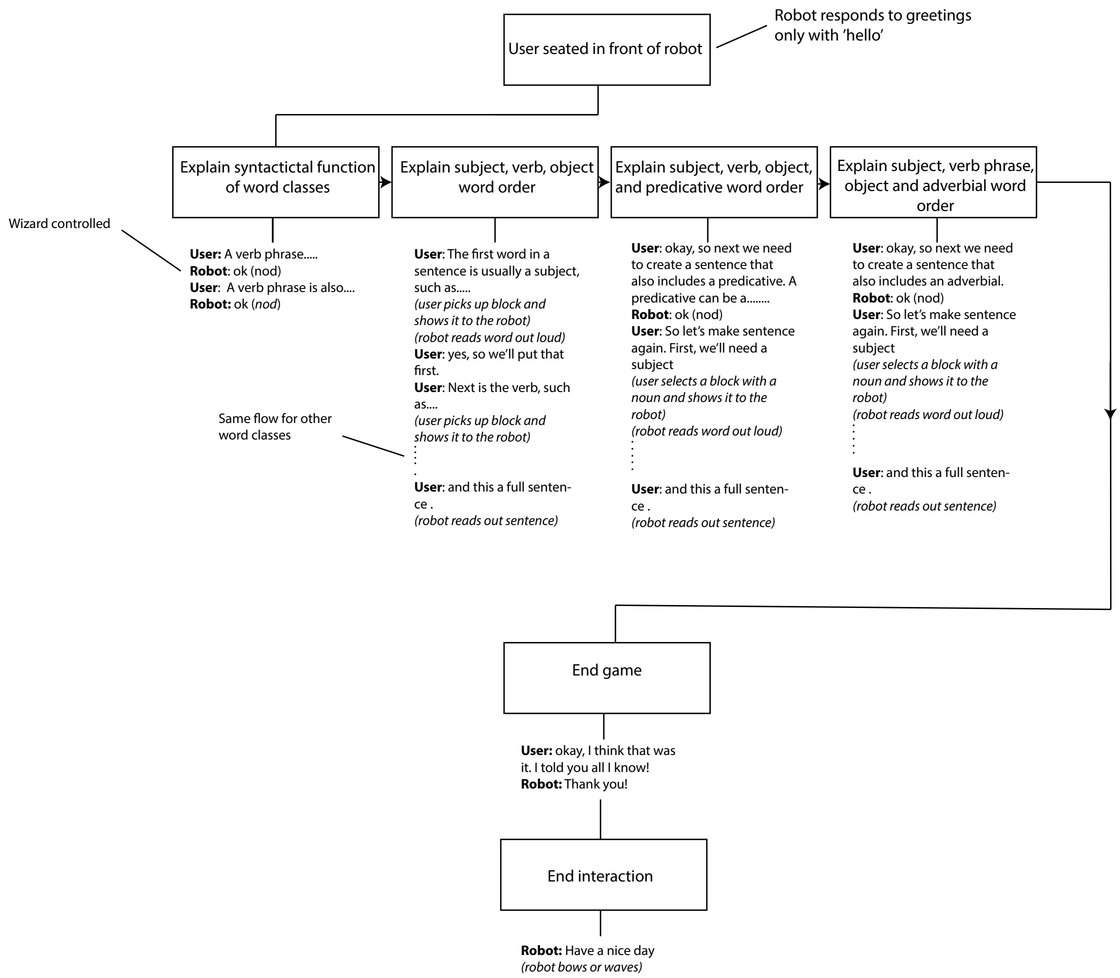


Figure 3.2: Interaction Flow

3.4.3 Lego mock-up

The word blocks consisted of 3 Lego Duplo blocks stacked vertically together. The blocks (put together) measured 6.5 cm by 6.5 cm. The blocks had printed a hamming marker on the one side (see figure 3.3), and a word on the opposite side. At this size, markers are recognizable by the software up to an approximate range of 1.5 meters. How markers and words are matched can be seen in the following subsection. Hamming markers are generated and recognized using the AR-Markers Python module[5].

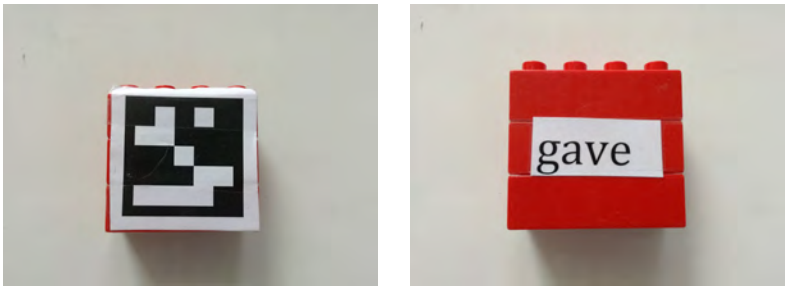


Figure 3.3: Lego word block

3.4.4 Technical setup

The interactions were carried out with a JD Humanoid, developed by EZ-Robots[6] (see figure 3.4). The robot is 31.8 centimeters tall with 16 degrees of freedom. The robot's speech was produced with MaryTTS (Schröder & Trouvain, 2003) with a UK English female voice.

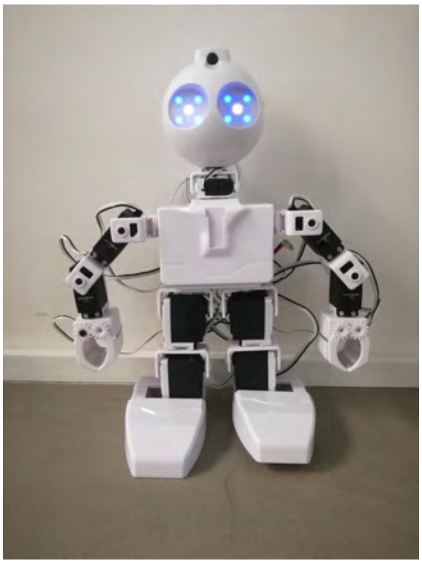


Figure 3.4: EZ Robot JD Humanoid

The robot was remotely controlled via three custom built python modules. For the autonomous nonverbal behaviors the robot made use of the tutor spotter program, originally developed for the iCub (Lohan et al., 2011), as well as two custom built Python modules, pyJD and SPY (developed together with the Robotics Lab at Heriot-Watt University). Both modules are Open Source and available for download on GitHub. The pyJD module [7] is a wrapper for YARP (Metta, Fitzpatrick, & Natale, 2006) and allows for control of the

[5] package available on GitHub https://github.com/DebVortex/python-ar-markers
[6] https://www.ez-robot.com/
[7] https://github.com/BrutusTT/pyJD

robot through YARP. The SPY module[8] was used to extract information on participants' gaze from a webcam. This information was then passed on to the tutor spotter program (Lohan et al., 2011), which then directly controlled the robot's gaze. Finally, the last module, called SlowWorm, allows the wizard to take control of the robot, time when the robot should speak, and to a certain extent also what it should say. See Figure 3.5 for an overview of what resources the wizard had at his/her disposal. The robot was able to read out words (or whole sentences) by scanning the hamming markers on the Lego blocks (see Figure 3.3). Each Lego block was encoded with one word. The SPY module was able to decode the hamming markers and pass the unencoded information on to the wizard control program (SlowWorm).

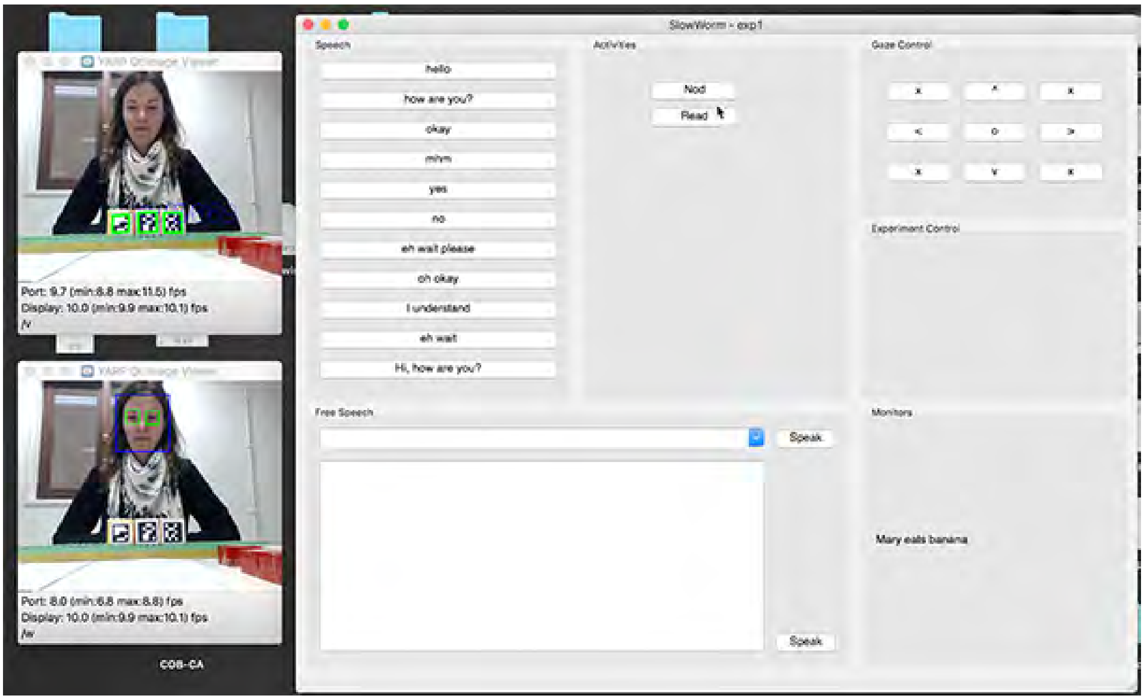


Figure 3.5: Control Interface

3.4.5 Subjective Measures

The questionnaire data was collected as subjective measures. The questionnaire consisted of two parts, a demographic part administered before the experiment took place, and a feedback part, in which participant rate and review their experience after the experiment. The demographic part of the questionnaire asked participants for their age, gender, academic training, field of study, and their previous experience with robots.

The feedback part of the questionnaire was divided into three separate parts, a part about the interaction; a part about the robot, and a part about the game participants are playing with the robot. The distinction between the three parts were purposefully made in such a way so that it would be possible to distinguish between participants' evaluations of the robot, the interaction and the game itself separately. The subjective measures chosen for this experiment probe participants' enjoyment and their perception of the robot's

[8] https://github.com/BrutusTT/spy

competence. Questionnaires for the interaction and for the game were asked as 7-point semantic differentials, while questions about the robot as asked on a 7-point likert scale.

The interaction was:

- *fun-boring*
- *long-short*
- *exciting-tiring*
- *engaging-boring*

The robot was:

- *intelligent*
- *friendly*
- *engaging*

The game was:

- *fun-boring*
- *dynamic-repetitive*
- *engaging-boring*
- *easy-difficult*
- *appropriate-not appropriate*

3.4.6 Objective Measures

Linguistic Analysis

All interactions were transcribed, coded and analyzed using the transcription software CLAN (MacWhinney & Wagner, 2010). The linguistic analysis is aimed at describing three aspects of participants' production, and uses a methodology developed by Fischer 2013, 2012. The three aspects are linguistic complexity, Interpersonal relationship, and signs of confusion. Linguistic complexity can be taken as evidence that speakers adjust their production to what they think their communication partner is able to understand. Signs of interpersonal relationship, for example the use of pronouns, can be taken as an indicator for the extent to which a participant considers the robot to be a social actor. How participants deal with interactional trouble, or whether they do at all, can likewise be taken as indicators for the extent to which participants treat the robot as a social actor. For linguistic complexity, the mean length of utterance (MLU), expository utterances, and the number of utterances per turn are measured. The MLU covers the average number tokens (words) in an utterance. A low MLU usually means less linguistically complex utterance. Expository utterances are utterances, which are explanatory in nature, and usually come in the form of "This is a...", "a verb functions as a....", or similar. As Fischer, Lohan, Saunders, et al. (2013) argue, high numbers of expository utterances may mean that one understands the robot is in need of explaining, but also that it is worth doing the explanations, i.e that it is treated seriously as an interaction partner. A high occurrence of expository utterances usually means that speakers reduce complexity, by explaining which objects of concepts they talk about. A high number of utterances per turn usually means that speakers present information in small 'bites', which in turn lowers the linguistic complexity. Interpersonal relationship includes the use of personal pronouns, reference to previous events or utterances, questions and tag questions. Signs of confusion include the number of times participants turn to the experimenter, the number of self- and other-repairs.

Prior to analysis, utterances that are clearly addressed to the experimenter are coded as "toEXP" on a subtier, so that they can be removed from subsequent analyses. Certain measures (mean length of utterance (MLU), number of utterances, utterances per turn, and personal pronouns) are calculated using CLAN's *freq* and *combo* commands. First, *combo* subsets the data to only include the speech of the participant (SP01). Second, the processed transcript is then piped to CLAN's *freq* and *mlt*, which provides measures for number of tokens (words), type (diversity) and type/token ratio, number of utterances and number of turns, as well as a word frequency list:

Combo +t*SP01: +t*ROB: +t*EXP: -s"toEXP" P1Elan.cha +d | mlt spause[9]

Interpersonal markers (questions, tags, and references to previous events) are coded manually and are subjected to a similar processing in CLAN:

Combo +t*SP01: +t%+s"expo" participant.cha +d | freq

Likewise, repairs (other-repairs, and self-repairs) are identified using a conversation analytical approach (Sacks et al., 1974). Utterances that include either form of repair are tagged and subjected to similar analysis as presented above.

Gaze Behavior

Gaze behavior is measured in duration (seconds) of the total duration of the interaction, in gaze occurrences per minute, and in the mean gaze length (per gaze action). This is done by manually coding the transcript produced in the previous section in ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Transcripts used for the linguistic analysis is reused for this analysis by converting the transcripts to the ELAN format by using CLAN's *chat2elan* function:

chat2elan +e.mov file.cha

3.4.7 Statistical Analysis

Both the subjective and the objective measures are analyzed using multiple linear regression. Predictor variables include the experimental condition, experience with robots and participant gender. These variables are known from the literature to have an effect on interaction with robots (see Chapter 1). To address these differences, participant gender and previous experience with robots are added as control variables in the regression models. Dependent variables include the twelve questionnaire items described above in the Methods section.

[9] Commands are colored in blue, switches colored in red

3.4.8 Participants

The experiment was initially pilot-tested on 15 participants at the University of Southern Denmark (Campus Odense). However, the experiment itself was carried out at the Vrieje Universiteit Brussels. 33 participants were recruited among students and staff. However, two participants chose not to complete the experiment, and the robot malfunctioned in three interactions. This leaves 28 interactions that are considered for analysis. Students come primarily from studies in mechanical engineering and psychology. There are slightly more women (60%) than men (40%), but they are evenly distributed across experimental conditions. None of the participants had previously interacted with the EZ-Robot. Participants had a mean age of 25.96 (SD=5.51). Participants were recruited from all levels of academia and include both first year bachelor students and professors.

3.5 Results

3.5.1 Questionnaire

Results of the questionnaire analysis reveal no differences between the three conditions in comparison of the ratings of the interaction. Participants rate the interaction as more *engaging* and *fun* in the *random gaze* condition compared to the other two conditions. However, these results do not reach statistical significance (see Appendix A). Generally, ratings in the two contingent conditions are very similar, and no significant differences between the two conditions are observed.

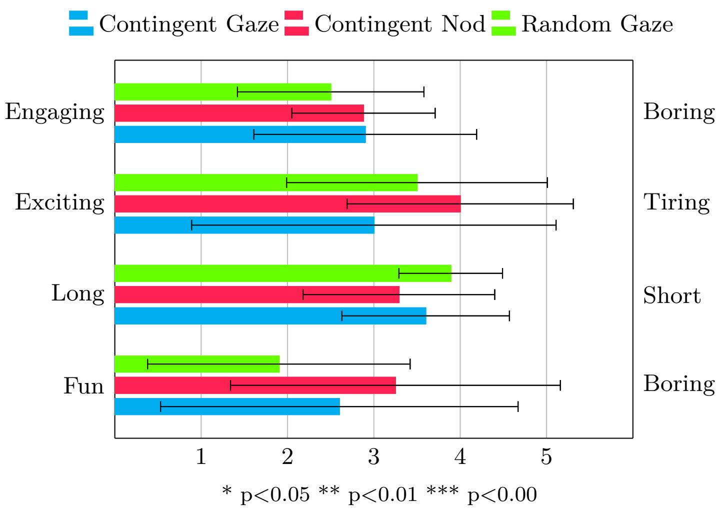


Figure 3.6: Questionnaire Responses for 'Interaction' Items

For ratings of the robot, participants rate the robot in the *contingent gaze* condition (B=1.08, se=0.44, p=0.02) as significantly more *friendly* than in the *random gaze* condition (see Figure 3.7).

Results of the ratings of the game show no differences between conditions (see Figure 3.8)[10].

[10] All regression tables are available in Appendix A

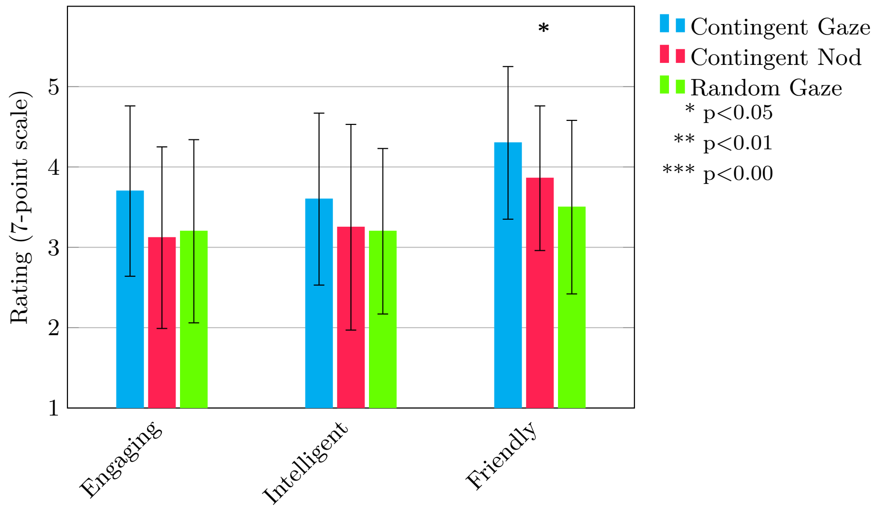


Figure 3.7: Questionnaire Responses for 'Robot' Items

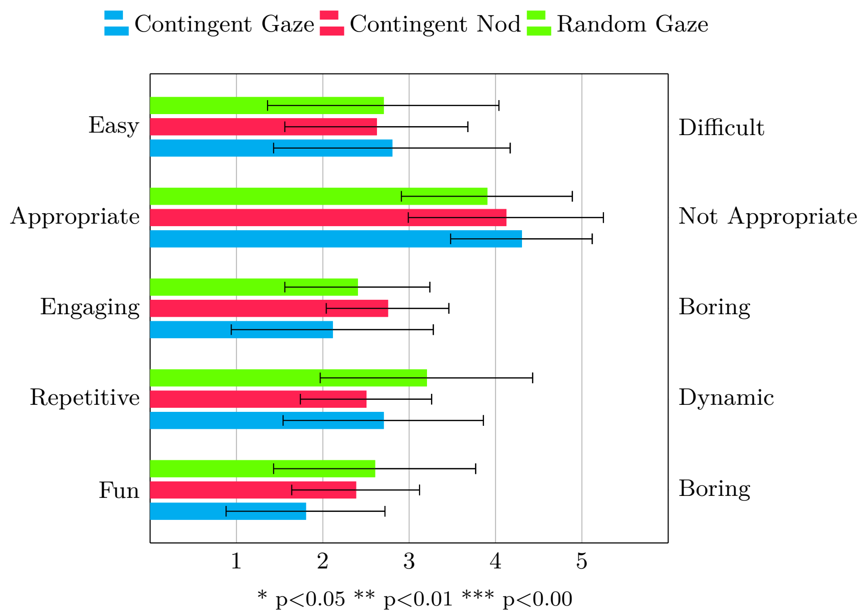


Figure 3.8: Questionnaire Responses for 'Game' Items

3.5.2 Linguistic Analysis

The analysis of participants' language use reveals subtle but significant differences between conditions in how they engage with the robot. Participants in the *contingent gaze* condition produce significantly more expository utterances in comparison to participants in the *random gaze* condition (B=-11.24, se=3.5, p=0.003), and marginally more in comparison to participants in the *contingent nod* condition (B=-8.88, se=4.4.5, p=0.053). Contrary to expectations, participants in the *contingent gaze* condition produce significantly fewer utterances per turn in comparison to participants in the *contingent nod* condition (B=1.29, se=0.52. p=0.02), and marginally fewer utterances per turn in comparison to the participants in the *random gaze* condition (B=0.86, se=0.42, p=0.051).

Participants in the *contingent gaze* condition produce significantly more markers of reference

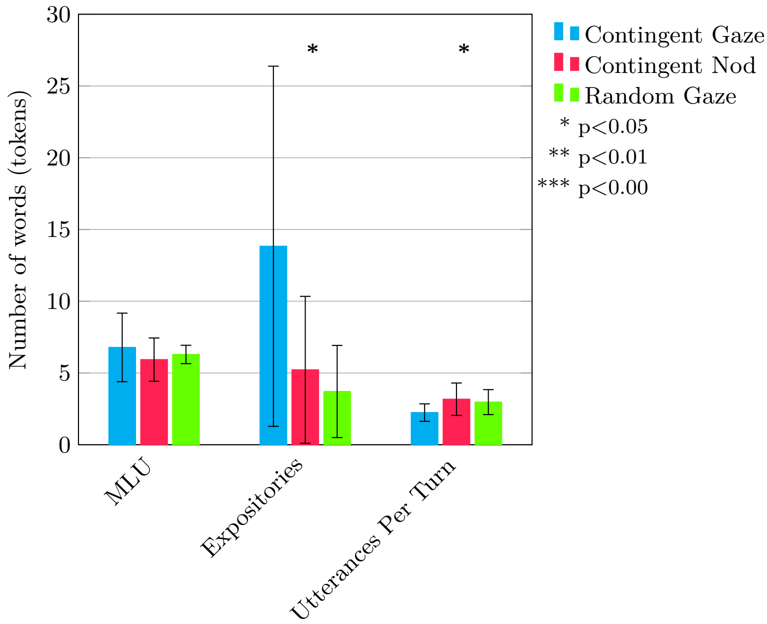


Figure 3.9: Linguistic Complexity

to previous actions and events than participants in the *random gaze* condition (B=-1.64, se=0.79, p=0.05), and also more markers in comparison to the *contingent nod* condition (B=-1.17, se=0.98, p=0.25), although the latter do not reach statistical significance. In addition, participants in the *contingent gaze* condition ask fewer, and not more questions in comparison to participants in the *random gaze* (B=4.73, se=2.59, p=0.08) and *contingent nod* (B=5.27, se=3.22, p=0.12) conditions. However, the differences are not statistically significant.

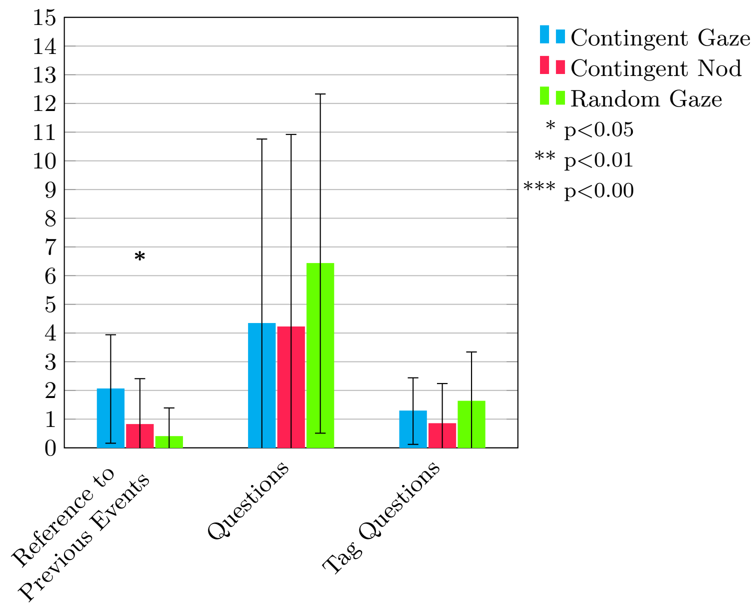


Figure 3.10: Interpersonality I

The analysis of the use of pronouns reveals a significant difference of first person singular pronouns (I, me) between the *contingent nod* condition (B=-16.86, se=5.27. p=0.004) and the *random gaze* condition. Likewise, there is a marginally significant difference of first person singular pronoun use (I, me) between the *contingent gaze* condition and the *random gaze* condition (B=8.92, se=4.78, p=0.07, see Figure 3.11. No difference in the use of first person plural pronouns (us, we) or in the use of second person pronouns (you) is found.

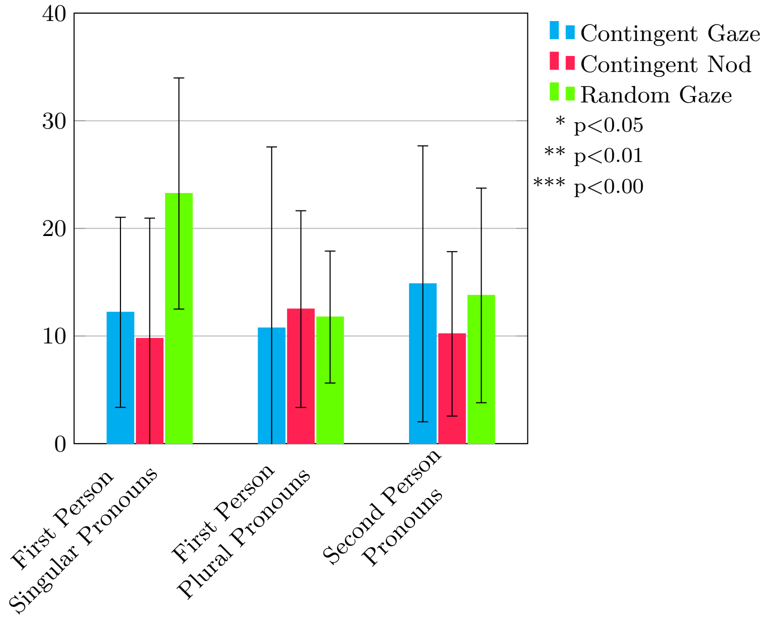


Figure 3.11: Interpersonality II - Personal Pronouns

Participants in the *contingent nod* condition turn to the experimenter significantly fewer times in comparison to participants in the *contingent gaze* (B=16.08, se=7.36, p=0.04), and *random gaze* (B=16.31, se=6.53, p=0.02) conditions. With regard to repairs, no significant differences between the conditions are found.

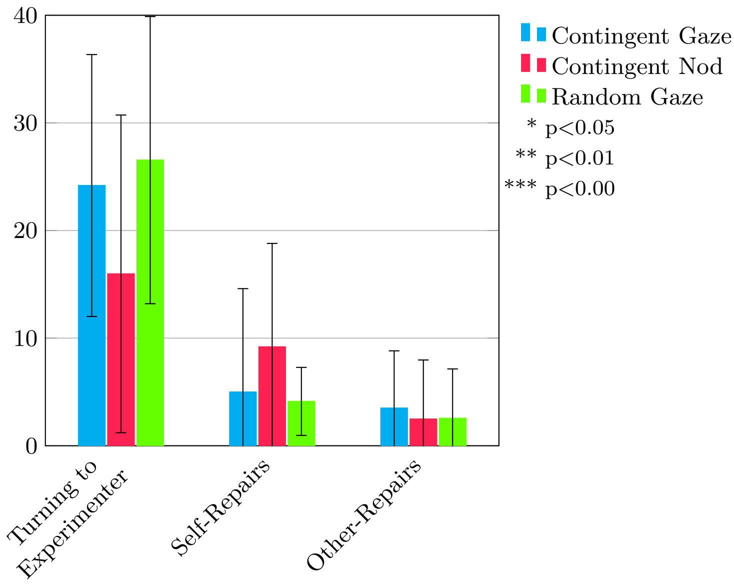


Figure 3.12: Signs of Confusion

3.5.3 Analysis of Gaze

None of the conditions reach the 30/70 ratio of human interaction (Argyle & Graham, 1976). Analysis of the number of participants' gaze actions reveals that participants in the *contingent nod* condition produce significantly fewer gaze actions per minute in comparison to participants in the *contingent gaze* condition (B=2.19, se=1.03, p=0.04), and participants in the *random gaze* condition (B=2.19, se=1.03, p=0.04). Conversely, participants in the *contingent nod* condition sustain their gazes for longer time than participants in the *contingent gaze* (B=-1.73, se=0.74, p=0.03) and participants in the *random gaze* (B=-1.6, se=0.74, p=0.4) conditions do (see Figure 3.13).

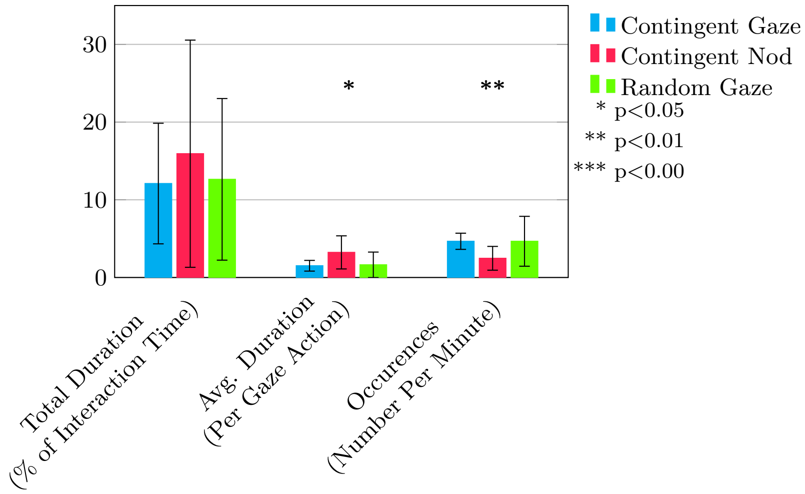


Figure 3.13: Analysis of Gaze

3.5.4 Results Overview

All significant differences reported above are summarized in Table 3.1 below.

	Contingent Gaze	Contingent Nod	Random Gaze
More expository utterances	X		
Fewer utterances per turn	X		
More references to previous events	X		
Fewer contacts to experimenter		X	
Fewer gaze actions		X	
Longer mean gaze duration		X	

Table 3.1: Summary of Significant Results

3.6 Discussion

3.6.1 Effects of Contingent Gaze

Although no differences between conditions in participants' subjective responses were found, several aspects of participants' behavior differ between conditions. One of the assumptions of the study is that people tailor their communicative behavior to their current communication partner. Previous work has shown that people do exactly this when interacting with a robot that uses contingent gaze (Fischer, Lohan, Nehaniv, & Lehmann, 2013; Fischer, Lohan, Saunders, et al., 2013)). Specifically, people reduce the linguistic complexity of their own production by splitting up their communication into smaller utterances, and by providing explanations for concepts through the use of expository utterances. Interestingly, participants in the current study produce expository utterances to a greater extent in the *contingent gaze* condition in comparison to the other two conditions. This result confirms results previously found with the same contingency spotter (Fischer, Lohan, Nehaniv, & Lehmann, 2013; Fischer, Lohan, Saunders, et al., 2013) and shows that the results also extend to a different robot morphology. Interestingly, the current study was not able to replicate the result (Fischer, Lohan, Nehaniv, & Lehmann, 2013) that participants design their turns over several utterances, thus producing more utterances per turn. In fact, the exact opposite was the case. That is, participants in the *contingent gaze* condition produced fewer rather than more utterances per turn and they assumed higher degrees of competence.

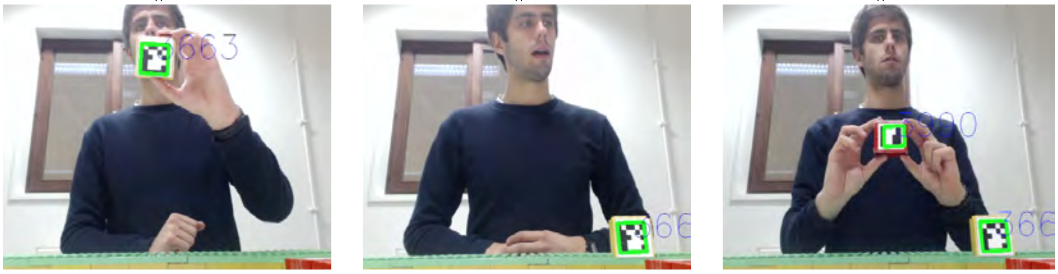
Establishing and Maintaining Common Ground

In order to better understand these seemingly contradictory results, the interactions were analyzed again. However, rather than looking at behavioral features that can be quantified, the analysis is based on a conversation analytical approach (Sacks et al., 1974). The qualitative analysis reveals two recurring themes that likely lead to the results reported above. Participants in the *contingent nod* condition and in the *random gaze* condition display an orientation to the importance that the robot is able to read the hamming markers on the Lego blocks. In addition, these participants do not plan the sentences they wish to construct ahead of time, but construct sentences incrementally one block at a time. During this process they produce several tokens that display an intention to keeping the turn. These two phenomena are demonstrated in Example 3.1 below.

The participant holds up a Lego block in front of the robot (#1) and asks whether it can `see` it (line 1). The robot reads out the word, which is understood as a demonstration of understanding and responded to by the participant by a positive assessment. The participant starts in line 5 with the discourse marker `so`, used here as a boundary marker for the coming talk. The participant then starts to look for candidate words among the Lego blocks, and while doing so produces a hesitation marker with a continuing intonation (line 8). This is followed by another hesitation marker (line 10) together with the last word the robot read (`jo:hn`), which is produced with a prolongation of the vowel and a continuing hesitation. As a result of these actions, he keeps the turn until he finds a second word block and holds it up (#2). Holding up a word block like this works as a turn yielding cue, which implicitly tells the robot to read the word block. After the robot reads out loud the two word blocks that are now visible, the participants responds with the assessment `goo:d` in line 13. The turn yielding cue of holding up the word block and the participant's positive assessment of the robot's behavior aptly display the participant's orientation as to what is important. It also shows an understanding of the robot's capabilities. It shows that the participant expects that the robot will respond in a certain manner to being shown a Lego block (namely, translating the code to a readable word and reading the word

```
01    *SP01:   dear robot can you see this?                          #1
02    *ROB:    john.
03    *SP01:   correct.
04             (.)
05    *SP01:   so i:'m gonna tell you something about john.
06 ((participant looks toward the stack of word blocks))
07             (2.5)
08    *SP01:   e:hm,
09             (1.8)
10    *SP01:   ehm jo:hn,                                            #2
11 ((participant holds up a word block))
11             (2.9)
12    *ROB:    john slices.
13    *SP01:   goo:d.                                                #3
```

#1 #2 #3



Example 3.1: Participant in the Random Gaze Condition

out loud). Using the next-turn-proof procedure, it becomes clear that this is what the participant expected the robot to do, as he assesses the robot's turn positively both times. This phenomenon is observed throughout the interaction, but also in other interactions, although primarily in participants who interacted with the robot in the contingent nod or the *random gaze* conditions. Likewise, the floor managing produced by the participant naturally prolongs his turn.

However, participants in the *contingent gaze* condition behave quite differently under similar circumstances. Rather than displaying an orientation to the robot's ability to read the word blocks, participants in the *contingent gaze* condition make other aspects of their communication with the robot relevant. Specifically, while they hold up word blocks as was seen in Example 3.1, they do not wait for the robot to read the word out loud, nor do they produce positive assessments of the robot's ability. Instead, they spend time explaining the grammatical concepts that each of the word blocks represents, thus making an implicit reference to the introductory phase, in which the concepts were first introduced to the robot. This phenomenon is demonstrated in Example 3.2.

The participant starts with the discourse marker `so`, which, as in the previous interaction, works here as a boundary marker for the next action. Thus, `second one` refers to the second sentence the participant is going to construct (having just completed the first one). By saying `second one`, the participant shows here that she considers previous as actions as common ground between them. This can be seen as evidence that he assigns epistemic competence to the robot. As also seen in Example 3.2 above, the participant holds up the word block while saying `john`, but already after the micro pause in line 1 she moves

```
01    *SP01:   so second one (.) john.                            #1
02 ((participant attaches the word block to the platform))
03             (1.8)
04    *SP01:   john.
05    *ROB:    john.
06    *SP01:   mm hm that's the subject of a sentence.            #2
07             (0.5)
08    *SP01:   then (.) eats.
09             (0.4)
10    *ROB:    john eats.
11    *SP01:   yes (.) to eat is a verb.                          #3
```

#1 #2 #3

Example 3.2: Participant in the Contingent Gaze Condition

the word to the platform. After she attaches the word block to the platform (line 2) she repeats the word, `john` (line 4). The robot reads out the word in line 5, but this is responded to only by a minimal response, which is followed by an exposition of the word's grammatical position in the sentence she is forming. She does this by tapping on the word block 'john', while already holding up the next word block (#2). The robot then reads out the two words now visible. Although this is responded to by the participant with a positive assessment (`yes`), it is within the same utterance followed up by another exposition of the grammatical position the word has in the sentence being formed. Thus, the participant here does not perform any work on the floor management, since it is not necessary. While she displays an orientation to the robot's ability to read the word blocks as a conditional relevant next action (as seen in lines 6 and 11), it is treated here as a cue to teach the robot something new, rather than to confirm that the robot is working within specified parameters, which is the case for Example 3.1. Again, this phenomenon is recurring not only in this particular interaction, but also in other interactions, in which participants interact with the robot in the *contingent gaze* condition. These two interaction formats are summarized in Figure 3.14.

The quantitative analysis suggested that participants reduce the linguistic complexity in their communication with the robot for some measures (expository utterances), while the opposite were true for other measures (number of utterances per turn). However, the qualitative analysis presented above suggests that the linguistic features investigated are used by participants in very different ways. Fischer, Lohan, Saunders, et al. (2013), Lohan et al. (2011) argue that contingent gaze allows people access to the robot's competences, and as a result thereof they adjust their communication accordingly. For participants in

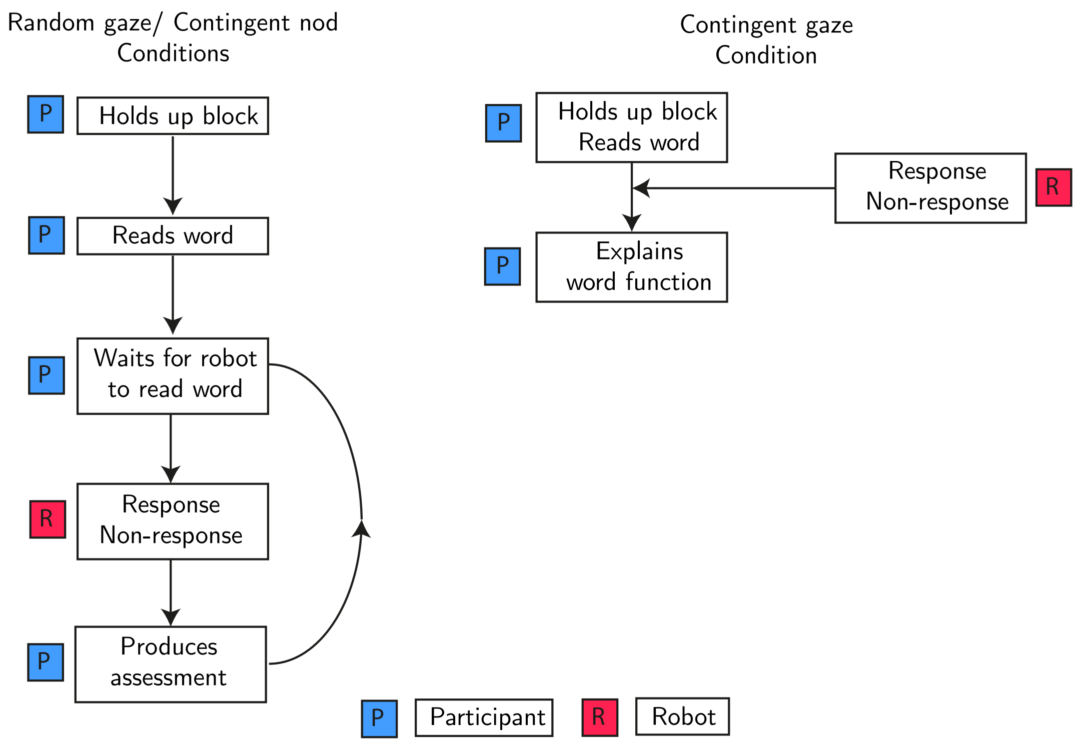


Figure 3.14: Interaction Formats

the random gaze and *contingent nod* conditions of the current study the gaze was not meaningful. Some evidence of this is found in participants' need to confirm what words are in use, and to asses the robot's performance in that confirmation. However, for participants in the *contingent gaze* condition the gaze was indeed meaningful and displays not only an awareness towards verbal actions (which is equal for all three conditions), but also for non-verbal actions. Therefore, participants did not need to confirm what the robot knows, but could move on to start teaching the robot. This conclusion is additionally supported by the number of references participants make to previous events. Here, participants in the *contingent gaze* condition not only implicitly refer to the introductory phase as seen in Example 3.2, but also make significantly more explicit references to previous actions and events. These references are realized in utterances such as *"this is similar to..."*, and *"as I said before"*. That is, participants in the *contingent gaze* condition establish early on in the interaction a common ground (Clark, 2002) through which the interaction unfolds and builds upon. Part of this common ground is everything that has occurred up to this point in the interaction, an assumption that participants in the other condition are likely not to make.

Contingency or Timing?

The results so far point in the direction that contingent gaze gives the impression of a responsive and situationally aware agent with which one can interact with in socially meaningful ways. As such this is a confirmation of previous work, which reach similar conclusions (e.g. Fischer, Lohan, Saunders, et al. (2013), Fischer et al. (2012), Lohan et al. (2011), Skantze et al. (2014)). In addition, it demonstrates that results obtained in previous work are stable even when adopted to a different robot morphology.

Furthermore, one of the aims of this study was to investigate the differences between contingency and timing. That is, the study addresses whether the positive aspects of contingent gaze, reported here and by others, can be attributed to the gaze action being contingent, or whether the gaze be replaced by any other contingent action, as indicated by the results reported by Mehlmann et al. (2014). The current study suggests that gaze alone (as for example the *random gaze* condition) does not bring forth the positive aspects discussed above, either. Likewise, replacing the gaze action with another contingent modality (as for example the *contingent nod* condition) does also not bring forth the positive aspect of contingent gaze either. Thus, contingent gaze is a social behavior that people recognize and which displays an awareness to human behavior in social interaction. This display directly influences how the interaction unfolds, what assumptions people make about their interaction partner, and which linguistic resources they put to use. However, people may be completely unaware of this influence, as suggested by the lack of differences between conditions for the subjective ratings. In conclusion, contingency is not merely a matter of producing ‘some action’ at the right time, but is an interaction governed by timing, conditional relevance, and modality.

3.6.2 Effects of Contingent Nods

The quantitative analysis revealed two interesting effects of contingent nods. First, participants in the *contingent nod* condition turn to the experimenter significantly fewer times than participants in either of the other two conditions. Second, participants in the *contingent nod* condition make fewer gaze action that last longer.

Social Actor or Not?

This seems to suggest that the nod has a disambiguation function. However, there is nothing in the qualitative or quantitative analyses that would indicate that this is indeed the case. An alternative explanation is that participants do not consider the to be robot a social actor in the *contingent nod* condition and thus are less worried about making mistakes. However, if this was the case, the number of self-repairs as well as other-repairs should be lower in comparison to the other conditions. This is not the case. In fact, participants in the *contingent nod* condition produce more self-repairs (see Figure 3.12) than participants in the other conditions, even though the difference is not significant. There is also no differences in the number of other-repairs. Nods can, however, work as a signal for affirmative action, urging participants to continue on with the ‘game’, a signal the robot does not produce in the other two conditions.

Participants in the *contingent nod* condition produce fewer gaze shifts than participants in other conditions, but maintain their gaze for a longer time. The lower frequency of gaze shifts in the *contingent nod* condition is likely linked to the robot’s own gaze behavior. In the *contingent nod* condition, the robot’s head is static for most of the interaction. Thus, it keeps its gaze fixated an approximate location of the participants’ faces (without tracking it).

One of the functions of gaze is information seeking (Argyle & Cook, 1976, p. 170), and gaze shifts indicate changes in the immediate environment that need to be attended to. However, since the robot's gaze almost never changes, there is no need for participants to monitor it as often. The robot in the *contingent gaze* and *random gaze* conditions differ in this respect. In these conditions the robot's gaze is constantly in a state of flux. The gaze produced in the *contingent gaze* condition is meaningful while the gaze produced in the *random gaze* condition does not seem to be relevant. What is relevant though, is that the robot's gaze is dynamic in these two conditions and is as such potentially relevant for the interaction.

In addition to producing fewer gaze shift, participants in the *contingent nod* condition maintain their gaze for longer periods of time. On the one hand, the *contingent nod* condition ensures that gaze is mutual as soon as the participant gazes at the robot. This can have both positive and negative consequences. Prolonged fixation of gaze is correlated with intimacy, but gazes that last too long can also seem threatening (Argyle & Cook, 1976). However, as mutual gaze is something that needs to be coordinated and negotiated (Kendon, 1967; Kendrick & Holler, 2017) by all parties in interaction, it can be argued that mutual gaze is never established in the *contingent nod* condition. That is, a coordination effort cannot be accomplished by doing nothing. From this argument follows that the robot is not treated as a social actor, but as an object equivalent to a painting or a statuette. This also means that participants do not need to be accountable for their gaze as the notions of threat and intimacy become irrelevant. However, had this been the case, this effect would also have been measurable in participants' linguistic productions and in their subjective responses to the robot and interaction, which it is not.

3.7 Conclusion

This study sought to investigate two related aspects of situatedness in HRI: contingency and timing. The study was able to replicate results from earlier studies on contingency. More importantly, the study was also able to show that while participants respond in very similar ways to the two types of contingent action, there are also significant differences between the two. In particular, the study showed that contingent gaze contributes to common ground and there are some indications that contingent nods reduce confusion about what to do. This seems to indicate that the two contingent actions perform distinct social functions in interaction. Contingent gaze can therefore not be replaced by other contingent actions, as hypothesized in the beginning of the chapter.

4. Study 2: The Discourse Record

4.1 Introduction

Situation awareness[1] has to do with what is going on in the environment as interaction unfolds, as well as making sense of things happening as they happen in the context of what has come before. As also discussed in Chapter 1, much work on situation awareness in HRI has to do with tele-operators' awareness of the physical environment when controlling a remote robot (Hedayati et al., 2018; Johnson et al., 2015; Scholtz et al., 2005; Zheng et al., 2011), or situation awareness of the people interacting with the robot (Begum, Huq, Wang, & Mihailidis, 2015).

Social interaction is highly dependent on people's ability to infer what their communication partners know and to adapt to the situation accordingly (Clark, 2002). Thus, communication partners make their actions and behaviors recognizable, observable (Sacks et al., 1974), and accountable (Garfinkel, 1967; Heritage, 1990). This in turn also means that people have an awareness of, and a moral responsibility to, what goes on in interaction with other people. Robots are usually not equipped with the same level of awareness, but are only aware to specific events and actions that are related to its primary function. However, as robots are increasingly designed for human social interaction, they need to be aware of behaviors that do not necessarily support specific goal driven task. This is potentially problematic since the range of human behavior is quite vast. In addition, people are not only aware of their communication partners' behaviors but also of the elements outside the local interaction context. Techniques that enable robots to track people's gaze (Lohan et al., 2011), face (Ishii et al., 2013), gesture (Suarez & Murphy, 2012), and body pose (McColl, Zhang, & Nejat, 2011) work well to display a robot's online ability to know where in a 3D space the human is located, and to some extent what activity he or she is currently engaged in (as also Chapter 3 demonstrated). For robots to be able to engage in meaningful social interaction, they need to be able to display an awareness to actions, states and events inside and outside the local interaction context, and do this using natural language, for example speech. Since human social interaction is contingent upon everything that has come before and sets the scene for what is expected to come next, a social robot will also need the ability to remember certain features of the interaction and be able to refer back to them at appropriate places during interaction with people. Here, research in social HRI needs to go beyond statistical information (such as game scores) and include a record of the verbal interaction that at the time of production does not necessarily seem

[1] Data collection for this experiment was primarily carried out by Master Student Anna Kryvous

salient or necessary for the interaction that follows.

4.1.1 Discourse Record in HRI

The part of the human memory system that has to do with the events, states and actions than can be explicitly recalled are referred to as the declarative memory (Squire & Zola, 1996). The declarative memory consists of two subsystems; the episodic memory and the semantic memory. Episodic memory stores specific events, states, and actions, as well as records of where they happened, when they happened, and the emotions felt at that point in time. The semantic memory stores facts, concepts and understandings of the world. It is generally understood that semantic memory draws on episodic memory for its contents (e.g. specific instances of learning how to ride a bicycle is stored in the episodic memory, but the ability to ride a bike is stored in the semantic memory). That is, as people gain more experience (episodic memory), they learn more about the world they are in (semantic memory). For robots, this relationship is usually slightly different. Not many HRI systems have an episodic memory, and those that do, do not store very many aspects of their communication; usually such systems save statistical data such as moves/wins/losses in a game (Ahmad, Mubin, & Orlando, 2017; Kipp & Kummert, 2016; Leite, Castellano, Pereira, Martinho, & Paiva, 2014) or performance on tests (Kasap & Magnenat-Thalmann, 2012; Leyzberg, Spaulding, & Scassellati, 2014). It is much less common to see episodic memory implemented in a system using speech, although there are exceptions (e.g. (Zhang, Zheng, & Thalmann, 2018)). Semantic memory in HRI is almost never derived from episodic memory, but hardcoded by human programmers. However, with recent developments in machine learning techniques and specifically deep learning, this will be subject to change in future social robot applications. It is imperative that robots be endowed with memory systems similar to humans, for robots to engage in rich social interactions (Christian, 2011).

To sum up, when it comes to access to episodic memory two major challenges (which each contain a range of smaller challenges) need to be solved. One is technical and concerns how a computational architecture can store and access data that is qualitative in nature. The other challenge is behavioral and concerns how the ability to display an awareness of current and past events, states and actions affects perception and the ensuing interaction. Thus, research into behavioral effects can shed light on which kinds of actions, states and events a robot should be able to remember and refer to. The current investigation attempts to address the latter challenge by experimentally studying the effects of a robot explicitly referencing to a current state of the world (commenting on the weather) and to an earlier user utterance. This is done in an experimental setting in which a robot instructs participants in how to build a Lego figure.

4.2 Literature Review

In the following, previous work on situation awareness in HRI is reviewed and discussed. As the current study investigates the conversational aspects of SA in HRI, special attention is paid to previous work that deals with conversation.

4.2.1 Indicators for Situation Awareness in Perceptual Basis

A large body of research investigates how robots display awareness through gaze cues, by either tracking a persons facial features (Ishii et al., 2013; Skantze et al., 2014; Zheng, Moon, Croft, & Meng, 2015) or a person's gaze (Baxter, Kennedy, Vollmer, de Greeff, & Belpaeme, 2014; Lohan et al., 2011; Pitsch et al., 2009). However, less attention has been paid to how robots can display situation awareness through natural language. Szafir and Mutlu (2012) investigate 'situation awareness' using an EEG monitor, with which the robot monitors and adjusts to participants' engagement levels. Kidd and Breazeal (2008) conducted a study in which a robot changes its greeting depending on the time of day. Likewise, Klamer, Allouch, and Heylen (2010) present a long-term study (10 days) in which a robot is installed in a user's home. While the overall linguistic capabilities of the robot were limited, it did have the ability to give participants a daily weather report. However, for both studies (Kidd & Breazeal, 2008; Klamer et al., 2010) it is unclear what the effect of the perceptual basis is. One study is purely exploratory and does not report results of this aspect of the robot's interactive capabilities (Klamer et al., 2010). The other study is controlled, but compares a robot with a tablet programmed with the same software and therefore does not report on any results directly related to the robot's situation awareness (Kidd & Breazeal, 2008).

A few studies have investigated how simulations of speaking autonomous cars can increase users' trust and acceptability as a communication partner, by displaying awareness to what happens outside the local interaction context (Sirkin, Fischer, Jensen, & Ju, 2015; 2016). These studies show that especially when the robot raises topics that are unexpected to users that users come to see the system as a serious interaction partner. However, the studies were not conducted in a controlled manner, and it is unclear whether these results also hold for embodied humanoid robots.

Thus, it is still relatively unclear how a robot can display its awareness to non-contextual events, states and actions through verbal conduct and what effects such a display have on the respective robot's communication partner.

4.2.2 Indicators for Awareness of the Discourse Record

Several researchers have called for more research into using memory in HRI studies (Baxter & Belpaeme, 2014; Castellano et al., 2008; Leite, Martinho, & Paiva, 2013). Some studies have already investigated how a robot's ability to store and recall interactional details affects how people perceive and interact with it. However, the results reported do not present a clear picture of the effects this has on people interacting with the robot. For example, there are several reports that show that when a robot displays the ability to access a memory, this leads to increased and sustained user engagement (Ahmad et al., 2017; Kasap & Magnenat-Thalmann, 2012; Leite et al., 2014) and to subjective ratings of the robot as more engaging (Kasap & Magnenat-Thalmann, 2012). However, not all studies are able to replicate this effect (Zhang et al., 2018), and comparisons with other types of manipulations show that adaptations to participants' perceived emotional state is more effective at increasing and sustaining engagement (Ahmad et al., 2017).

In addition, manipulations are not always tested with a control condition (Leite et al., 2014). Other studies have found that robots with a memory capability increase what authors refer to as 'positive affect' (Campos & Paiva, 2010; Leite, Pereira, & Lehman, 2017), anthropomorphization (Kipp & Kummert, 2016) or 'appropriateness' in dialog (Zhang et al., 2018). Finally, access to a memory module has also been shown to increase participants' learning (Leyzberg et al., 2014).

Several systems are already able to remember events and to use this information later in the current, or in subsequent interactions. For example, in an early study, Kanda, Hirano, Eaton, and Ishiguro (2004) implement a robot in a school. The robot is able to recognize children via an RFID chip. In another study, Kanda, Sato, Saiwaki, and Ishiguro (2007) deploy the Robovie in a school setting. The robot stores who it has interacted with, for how long and who else was present in the interaction. In a similar study using the same robot, Kanda, Shiomi, Miyashita, Ishiguro, and Hagita (2009) implement an identification system into their robot that makes the robot able to recognize persons it has met before and to make references to what that person has told the robot. Other works use the human face as identifier (Hanheide, Lang, & Sagerer, 2008; Kanda et al., 2009; Mykoniatis et al., 2013; Zhang et al., 2018). That is, facial features from interaction partners are saved in an initial encounter, and a robot is able to recognize that same person in subsequent encounters, using facial recognition. Common for these studies is that they are explorative in nature, and while the works are both relevant and novel, it is difficult too see exactly how the robots' ability to display awareness to the dialog history affects communication partners' perception of and interaction with it. Petit, Fischer, and Demiris (2016) present a technical framework that is able to store and retrieve episodic and semantic memory items and implement the framework on the iCub, Nao, and Baxter robots. Other ways to identify a person include extracting information from online social networks (Mavridis et al., 2010), and in some studies identification of users is done by the research team, so that participant identification is hardcoded prior to interaction (Leyzberg et al., 2014)

Other works are less concerned with remembering partners between interactions, but focus on keeping track of the local interaction history. Here, special attention is paid to the formal elements of interactions. 'Formal elements' are understood here as elements that advance the activity the interaction partners are currently engaged in. For example, making a move in a game or completing a specific task both constitute such elements. Several works have implemented this type of memory. For example, Leite et al. (2014) present a study in which children play chess with the iCat robot several times over the course of a semester. The robot is able to remember the outcomes of past games and to track, remember and recall all of the moves of the current game. Several other works implement memory in similar ways (Ahmad et al., 2017; Castellano, Pereira, Leite, Paiva, & McOwan, 2009; Kidd & Breazeal, 2008; Kipp & Kummert, 2016; Leyzberg et al., 2014).

A different approach to develop rich social interactions is to 'invent' a back story for a robot, for example, by giving the robot a name and a history, which it then is able to refer to. In this way, a robot is able to refer to events, people or objects in its fictional past in interactions with people (Simmons et al., 2011). Likewise, Swift-Spong, Wen, Spruijt-Metz, and Matarić (2016) compared fictional with non-fictional back stories, which the robot referred to at specific points during the interaction. Their results show however no differences in participant rating of the robot or their engagement with it.

Only few studies extract interaction partners' linguistic productions (i.e. speech) as features in a memory system. For example, Kasap and Magnenat-Thalmann (2012) present a system implemented in a robotic tutor. Here the communication between human and robot is verbal, but the system is only able tp store and recall responses produced by the human to test questions. Thus, the system does not go beyond the discourse record and ignores aspects of the human tutee's verbal contributions to interaction (as understood by Clark (2002)). Kanda et al. (2009) present a robot in a shopping mall that has the ability to remember features of talk that go beyond the discourse record (e.g. remembering that a particular person likes ice cream). However, since this functionality is not tested in any controlled way, it is unclear what its effects are. Zhang et al. (2018) advance work on memory in HRI even further by integrating speech, gestures, facial features and head movements in a module that is implemented on a virtual (on-screen) character. They show that participants find the virtual agent to be more appropriate and more useful with the memory module than without. However, they find no differences in user engagement or in responsiveness. In addition, the module is not tested on an embodied agent, such as a robot.

Thus, it still relatively unclear how currently a social robot's display of situation awareness, here in the form of access to memory of the discourse record, affects user engagement and participants' understanding of a robot's situation awareness.

4.3 Method

4.3.1 Experimental Conditions

The experiment featured two experimental conditions. In one condition, the *high-aware condition*, the robot commented on the weather in the beginning of the interaction. During the introduction, the robot asked participants whether he or she liked to play with Lego. Towards the end of the interaction the robot recalled the participant's response. In particular, the robot asked participants whether he or she thought this activity was fun (in case they said they do like to play with Legos) or if it was fun despite their previous negative stance toward Lego. Prior to the experiment, the human wizard defined states for the robot, such as what the current weather is like and the robot's level of awareness (high or low). In the *high-aware* condition when the weather was good the robot said:

> *What a nice day to play, next time we should do it outdoors.*

Whereas for the same condition, when the weather was bad, it said:

> *Such a bad weather today, it's nice to stay inside and play.*

In the *low-aware* condition this utterance was skipped completely. Next, regardless of condition the robot asked:

> *Do you like playing with Lego bricks?*

In the *high-aware* condition the human wizard noted down whether the participants expressed agreement, and if that was the case, the robot said:

> *It should be interesting for you then.*

Regardless of the user response the, robot said this also in the *low-aware* condition. In the *high-aware* condition when the wizards identifies a response that expressed disagreement, or if the participant simply did not respond, the robot said:

> *Then it's good that you're here. Maybe you get to enjoy playing with it.*

At the end of the interaction, the robot in the *high-aware* condition recalled what the participant previously had expressed about his or her stance towards Legos. If participants expressed agreement, the robot says:

> *You said before that you liked playing with Legos. Did you enjoy it?*

Whereas had he or she expressed disagreement or not responded to the question, the robot said:

> *You said before that you didn't like playing with Legos. Did you enjoy it?*

In the *low-aware* condition, the robot make no reference to participants' previous response, but said instead:

> *Playing with Legos is a good exercise for the brain. Did you enjoy it?*

4.3.2 Participants

52 participants were recruited from the University of Southern Denmark, campus Sønderborg. Mean age was 23.6 (SD=3.6). Only 26.9% of the participants were women and these were not perfectly distributed between the conditions. Thus, nine women interacted with the robot in the *high-aware* condition, while five interacted with the robot in the *low-aware* condition. Participants are ethnically very diverse; most of the participants come from Denmark or Germany, while other participants come from Australia, Bulgaria, China, Croatia, Iceland, India, Latvia, Lithuania, Moldova, Norway, Pakistan, Poland, Romania, Slovakia, and Spain. 42.3% of the participants have had experience with robots before, but participants with or without previous experience are distributed equally between the conditions.

4.3.3 Robot and Software

The robot used for the experiment was the EZ-Robot Humanoid JD identical to the robot used in Chapter 3, but with several modifications to its programming. Rather than using the Contingency Spotter as in Chapter 3, I programmed new behaviors for the robot using the EZ Builder application and the EZ-Robots Software Development Kit (SDK), which are detailed below.

Vision and Gaze System

The robot's internal VGA camera tracked the relative location of participants' faces from the center of the 2D video stream. These data were then coupled with the robot's vertical and horizontal head motor control to the effect that the robot always gazed towards participants. This functionality, which relies on the OpenCV library (Bradski, 2000), is turned on in both conditions. This functionality is referred to as 'face tracking'. This tracking is different from the contingent gaze tracking presented in Chapter 3. The system presented in Chapter 3 tracks the eyes of the participant and their projected gaze, while the current method uses facial recognition to track the face. However, probably due to the robot's relatively low degrees of freedom the two implemented methods produce visually very similar behaviors.

Dialog Management

In order to produce speech, the robot accessed the IBM Watson Text-To-Speech service, which produces speech on-the-fly[2]. All speech and nonverbal actions were pre-scripted and were to some extent timed by the wizard. This also applies to the feedback the robot supplied to participants as they progressed through the assembly. Feedback options available to the wizard were elicited by running pilot studies of the experiment, first with human participants, in which one plays the 'robot' and later also in a setup with a robot, similar to what is presented here[3].

4.3.4 Wizard-of-Oz Module

I developed an extra module as a plugin for the EZ-Builder program through which the wizard can easily control the robot (see figure 4.1). The module, written in C# with WinForms and the EZ-B SDK, consists of two types of elements, which manipulate the robot's state in various ways. One type sets explicit states for the robot, for example, the weather. This was done through a simple 'set' command together with an if-statement:

```
private void ddWeather_SelectedIndexChanged(object sender, EventArgs e)
{
if (ddWeather.Text == "Great")
{
EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "great");
}

if (ddWeather.Text == "Bad")
{
EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "bad");
}
}
```

Source Code 4.1: Weather Control

The other type of element is a single button called *intervention* (see figure 4.1). This button was used during the experiment by the wizard to time certain events, for example to

[2] https://www.ibm.com/watson/services/text-to-speech/)

[3] This work was done by Anna Kryvous, as part of her MA Thesis

continue the interaction after a participant response. The button was colored red to indicate to the wizard that the interaction was stalled until he or she pressed the intervention button, after which it would return to its basic gray color. This was done on the plugin side by incrementing a counter by one every time the button was clicked, set the variable in EZ-Builder with the same value as the counter, and return the original color to the button:

```
private void btnInterv_Click(object sender, EventArgs e)
{
intCount++;
EZ_Builder.Scripting.VariableManager.SetVariable("$intervention", intCount)
    ;
EZ_Builder.Invokers.SetBackColor(bntInterv, Color.LightGray);

}
```

Source Code 4.2: Intervention Button Code

On the EZ-Builder side, the script waited for a change in the variable *$intervention* before proceeding.

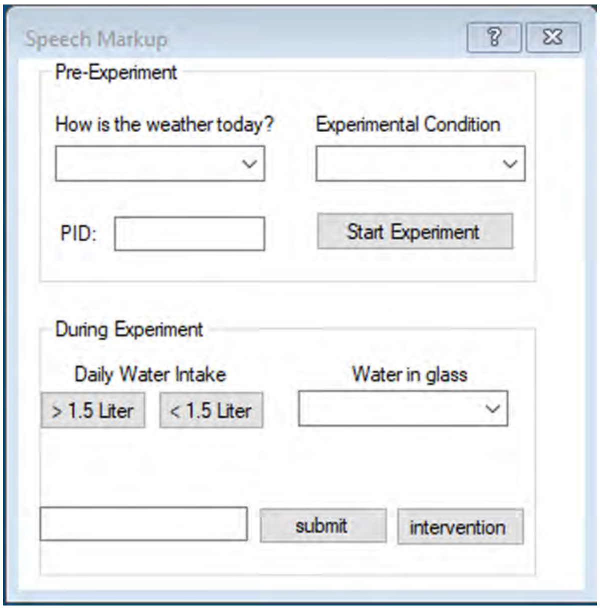


Figure 4.1: Wizard-of-Oz Module

The entire source code for the plugin can be seen in Appendix B.1.

4.3.5 Assembly Task

4.3.6 Analysis

Questionnaire responses are analyzed using linear multiple regression. Predictor variables include the experimental condition, experience with robots, and gender.

Subjective Measures

Prior to the experiment, participants completed a demographic questionnaire their eliciting info about participants' age, sex and previous experience with robots. After the experiment,

Figure 4.2: Assembled Lego Figure

participants were asked to rate the robot on nine different traits on a 7-point Likert scale as well as a single question on a 5-point likert scale that functions as a manipulation check. Participants were asked to rate the extent to which they think the robot is *social*, *interactive*, *reliable*, *competent*, *intelligent*, *knowledgeable*, *entertaining*, *boring*, and *engaging*. For the manipulation check, participants were asked about the extent to which they believe the robot was aware of their actions. These questions were designed to elicit participants' perceptions of the robot's situational awareness, sociability, and competence.

4.4 Results

4.4.1 Subjective Measures

Manipulation Check

For the manipulation check, participants were asked the extent (on a 5-point scale) to which they thought the robot was *aware* of their actions. Generally, participants in both conditions thought the robot was very much aware of their actions (see Figure 4.3). However, participants in the *high-aware* condition rate the robot as significantly (B=0.36, SE=0.14, p=0.01) more aware than participants in the *low-aware* condition. Thus, results show that the manipulation is indeed effective.

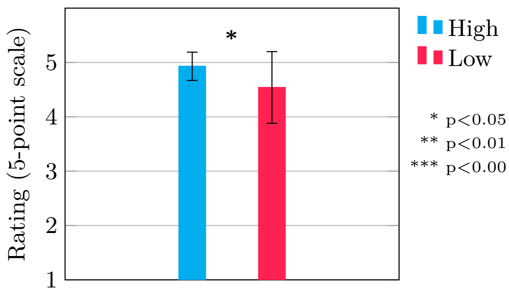


Figure 4.3: Manipulation Check

Questionnaire

Regarding participants' perceptions of the robot's competence, the analysis yields no significant differences. Likewise, neither gender nor previous experience with robots had any impact.

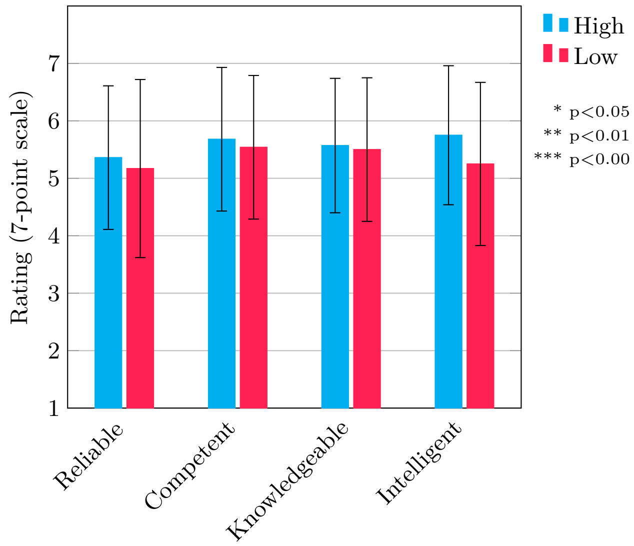


Figure 4.4: Evaluation of Competence

Participants rated the robot in the *high-aware* condition as significantly more social (B=0.87, SE=0.43, p=0.05) and more interactive (B=0.63, SE=0.30, p=0.04) than in the *low-aware* condition (see Figure 4.4).

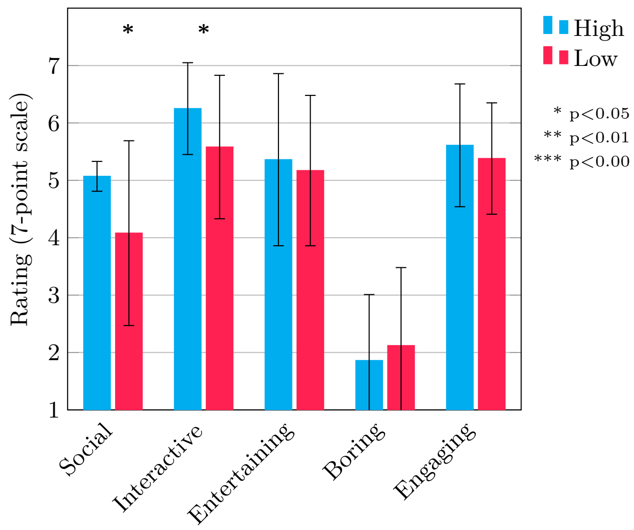


Figure 4.5: Evaluation of Social Features

Interpersonal Differences

The visualization of the regression model in Figure 4.6 shows the significant effects of the experimental condition reported on above, as well as significant effects of the dependent variables 'previous experience with robots' and participant gender. Experience with robots influences the participants' perception of the robot only to a very small degree. Specifically, participants who have previous experience with robots perceive the robot as less *intelligent* than participants who do not. However, this difference is only marginally significant (B=-0.69, SE=0.38, p=0.07). Participant gender influences the degree to which participants perceive the robot to be *boring* and *engaging*. Specifically, men perceive the robot to be significantly more *boring* (B=0.87, SE=0.40, p=0.03) and significantly less *engaging* (B=-0.87, SE=0.32, p=0.009), regardless of condition. Neither experience with robots nor participant gender interact with the experimental condition in any of the items in the questionnaire.

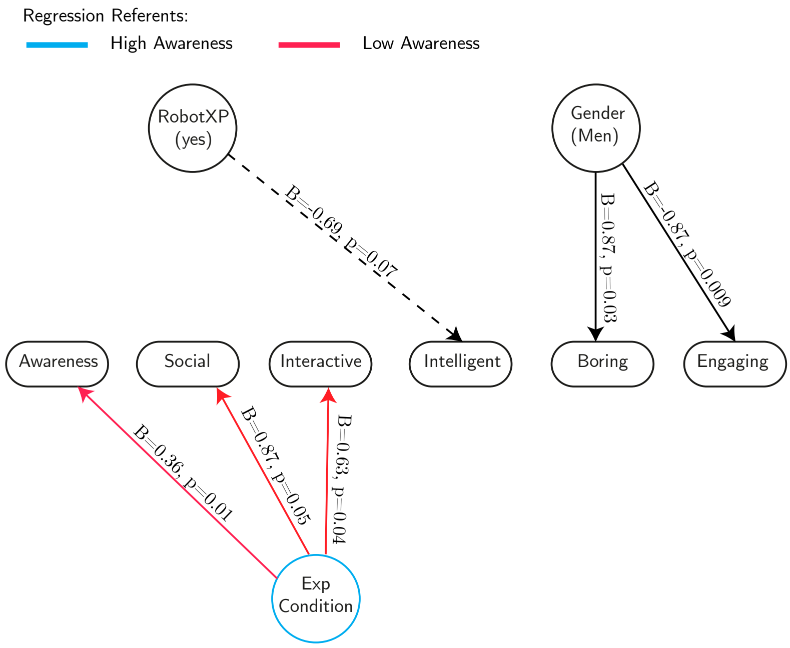


Figure 4.6: Regression Model

4.5 Discussion

The aim of this study was to find out how displays of situation awareness through attention to one aspect of common ground, namely the perceptual basis, and to (a limited) dialog history affects participants' perception of a social robot. The study found that even small adjustments in the robot's behavior lead to significant increases in participants' perception of the robot. Specifically, the study found that the robot was perceived to be more aware, more social, and more interactive with the awareness manipulations than without. This should be seen in the light that the robot already displays an attention to other contextual features, such as where participant were (face-tracking), how participants performed the assembly (feedback), and participants' verbal behavior (to some extent). That is, in both conditions the robot tracks participants' faces, responds contingently to utterances, and coordinates its speech with gestures. With this in mind, it is surprising that participants rate the robot as significantly more aware (although the difference between means is low) when interacting with the robot in the high-aware condition. Likewise, the robot is quite interactive through its verbal conduct, but again, using only two small manipulations, participants perceive the robot as significantly more interactive. However, other affective measures such as engagement and the degree to which the robot was entertaining or boring did not differ between conditions. This is in contrast to other work that reports positive effects on, for example, engagement, when a robot displays an awareness to the dialog history (Ahmad et al., 2017; Kasap & Magnenat-Thalmann, 2012; Leite et al., 2014). However, affective measures such as engagement are measured in very different ways. Some studies (Campos & Paiva, 2010; Kasap & Magnenat-Thalmann, 2012; Leite et al., 2014; Leite et al., 2017) measure affective factors using questionnaires, as was also the case for the current study. However, measurements are not always compared with a control condition, but rather analyzed over time using a repeated-measures approach. Thus, a measure is reported as "sustained" when no difference over time is found. Other works (Ahmad et al., 2017; Castellano et al., 2009) use gaze, facial features, and gestures to approximate participants' affective responses and do show that engagement and other affective factors are influenced by the robot's displays of awareness of the dialog history. Another explanation could be that affective factors are very complex cognitive processes that are difficult to formalize (Lemaignan, Garcia, Jacq, & Dillenbourg, 2016).

4.5.1 Effects of Awareness of the Discourse Record the Perceptual Basis

Much of the previous work on dialog history focuses on statistical data in game-like scenarios (Ahmad et al., 2017; Kipp & Kummert, 2016; Leite et al., 2014), on performance (Kasap & Magnenat-Thalmann, 2012; Leyzberg et al., 2014), or on previous communication partners (Hanheide et al., 2008; Kanda et al., 2009; Mykoniatis et al., 2013; Zhang et al., 2018). However, while many of these studies focus on affective and social features in the evaluation of these systems, only very few implement awareness to the dialog history with a basis in social, rather than statistical information. For those few that do (for example Kanda et al. (2009), Zhang et al. (2018)) it is unclear, due the experimental setup of the studies, how this awareness directly affects perception and interaction. In addition, some works show that robots should use some information (for example game statistics) only sparingly as they can influence participants' perception negatively (Kipp & Kummert, 2016). Awareness to the perceptual basis is usually considered a positive trait (Kidd & Breazeal, 2008; Klamer et al., 2010), and surprises people as they discover this awareness (Sirkin et al., 2015; 2016). However, not much work has been done to document what the perceived effects of this type of manipulation are. The current study bridges part of this gap by demonstrating how even a small manipulation of a robot's social awareness has significant effects on participants' perception of the robot's awareness, sociality, and interactivity. These results encourage further studies into how robots can influence people's perception of robots along the social dimension using social elements from the dialog history, and the perceptual basis in their interaction.

5. Study 3: The Perceptual Basis, Face-tracking, & Incrementality

5.1 Introduction

Chapter 3 showed[1] that contingent gaze helps to establish and maintain common ground between human and robot. Likewise, Chapter 4 showed that a robot's awareness to the perceptual basis can positively affect the extent to which participants perceive the robot to be interactive and social. The current study builds on these results and investigates different displays of the perceptual basis, and different displays of awareness to participants' verbal and non-verbal actions.

As established in Chapter 1, common ground can be established and updated in a number of ways. One way is to signal an awareness to the perceptual basis. The literature review for Chapter 4 concluded that research on the effects of displays of awareness to the perceptual basis is very limited, in an HRI context. Chapter 4 contributed to this research. Specifically, the robot displayed an awareness to the perceptual basis by commenting on the current weather situation. The study suggests that displays of awareness to the perceptual basis affect the extent to which participants perceive the robot as social and interactive. However, the study did not investigate any behavioral effects of these displays. It is thus still relatively unknown how displays of awareness to the perceptual basis affect interaction directly.

Another way to establish and update common ground is through gaze. Results from Chapter 3 suggest that contingent gaze positively contributes to the establishment of common ground. However, contingent gaze is by far the not the only gaze mechanism employed by people in interaction. In addition, contingent gaze may be less informative when the participants in interaction are not handling any tangible objects. An alternative gaze mechanism was implemented in Chapter 4, in which the robot continually follows a participants face by using a head pose estimation algorithm. While the mechanism did not enter into the experimental design many participants did report that they felt that the robot was aware of their actions. Similar to Chapter 4 many HRI studies implement face tracking, but surprisingly few studies investigate how the mechanism affect perception and interaction.

A third way through which to establish and update common ground is through incremental feedback. Incremental feedback should here been understood in the sense that information,

[1]The data collection for the current study was carried out together with MA student Nadine Petersen

in the form of speech, is presented in small chunks, with a basis in what is currently going on in the interaction, and what is needed to accomplish a specific action. As such, incremental feedback signals an awareness to the interactional context and gives an indication of what can be considered part of the common ground. Previous work on incrementality shows some inconsistencies in how systems providing incremental feedback are evaluated (see literature review below). Furthermore, incremental feedback is usually (when at all) compared to non-incremental feedback, but is rarely compared against other displays of awareness in interaction.

These three types of signals (awareness of the perceptual basis, face-tracking, and incrementality) of common ground are investigated in an experimental study in order to better understand how they affect perception and behavior of participants in interaction with a robot. The aim of this research is to understand the relative contributions of each of these signals to participants' perception and behavior.

5.2 Literature Review

The current study investigates two aspects of common ground, namely the perceptual basis and the actional basis. The actional basis is implemented in two different ways; as incremental feedback and face-tracking.

5.2.1 Incrementality

Incremental speech processing enables users to have online access to what contextual information a robot is attending to. There is already some work on implementing incremental speech processing in robots. For example Manuvinakurike, Paetzel, Qu, Schlangen, and DeVault (2016) assign utterances into 1 of 18 dialogue acts in their dialogue segmentation system, based on word-for-word processing of the speech input. However, the system is not tested in a live human-robot interaction scenario. Similarly, Carlmeyer, Schlangen, and Wrede (2014) present a dialogue system for use in HRI in which users can provide feedback and correction to the robot. However, the work is only presented as a proof-of-concept and not tested experimentally in interaction.

Other work shows that incremental speech can decrease response time since a system will begin production before it has stopped processing information relevant to that production (Skantze & Hjalmarsson, 2010; 2013). Kennington et al. (2014) implement incremental speech in a dialog system for a car simulator. Their study shows that participants interacting with the incremental system perform better at driving-related tasks than participants who interacted with a non-incremental speech system. Ghigi, Eskenazi, Torres, and Lee (2014) implement incremental speech processing in an information retrieval system. Their study shows that although dialogues become longer, the success rate is higher in the incremental condition. In other words, people do to a greater extent get the information they request and experience fewer problems. Chromik, Carlmeyer, and Wrede (2017) show that people interacting with a robot in an object-fetching task perform better when instructions are given incrementally than when given all at once. Thus, participants in the incremental condition we better able to find the correct items.

Studies also show differences in perception between incremental and non-incremental feedback systems. Generally, incremental speech systems are perceived as more polite and efficient (Baumann & Lindner, 2015; Skantze & Hjalmarsson, 2013), more natural (Buschmeier, Baumann, Dosch, Kopp, & Schlangen, 2012), as well as more responsive, enjoyable, and attentive (Tsai, Baumann, Pecune, & Casell, 2018) than non-incremental systems are. However, there is also evidence of the contrary. For example, while Baumann and Lindner (2015) report their simulated robot to be perceived as more polite and natural, Chromik et al. (2017) report that their simulated robot is rated as less natural when using incremental speech, and Carlmeyer, Schlangen, and Wrede (2016b) report that their simulated robot is rated as less likable when using incremental speech. de Kok et al. (2015) present a virtual coach that provides online feedback as participants do exercises, such as squats. Here, the feedback comes in the format of "watch your kneck", and "go a little deeper" as the system detects errors in participants' behavior. Evaluations of the system shows that incremental instructions were correlated with intelligence, helpfulness, responsivity, humanlikeness and clearly, but that the robot is also perceived as tiring. The behavior was generated on the basis of analyses of a corpus of interactions between an exercise coach and experiment participants (Hough, de Kok, Schlangen, & Kopp, 2015).

To sum up, research on incremental speech processing in HRI indicates that incrementality can contribute to efficiency, and there is some evidence that suggests that participants perceive the robot more positively, although there is also evidence of the contrary. It is also clear from the current review that incremental speech interfaces are only rarely used on embodied robotic systems, but rather on virtual avatars. It is therefore an open question how many of these findings also apply to embodied HRI.

5.2.2 Face-tracking

A common method in tracking the face of a person in HRI is to rely on head pose estimation. For example, using this method, Lemaignan et al. (2016) measure children's engagement with a robot. Other works use similar methods to measure engagement (Anzalone, Boucenna, Ivaldi, & Chetouani, 2015; Castellano et al., 2013), interaction quality (Baur, Damian, Lingenfelser, Wagner, & André, 2013), and to signal attention and awareness (Bohus, Saw, & Horvitz, 2014). However, for most of these studies, the robots do no change their own head pose, but rather use the information of a user's head pose as one out of several components in a computational model that hypothesizes about what a user is currently engaged in (or not). Face-tracking mechanisms are generally considered 'positive traits' (Asselborn, Johal, & Dillenbourg, 2017), but not much work investigates exactly how face-tracking contributes interactionally and what effects this has on how the robot is perceived. For example, several studies implement face tracking in their robots, but do not explicitly test its effects (Anastasiou, Jokinen, & Wilcock, 2013; Andrist et al., 2014; Yamazaki et al., 2009).

Only few studies report findings of experimentally tested effects of a robot tracking a user's head pose. For example, Riek et al. (2010) did not find any differences between groups of participants that interacted with a robot producing full facial mimicry, nodding, or no mimicry at all. In an earlier study, Wang, Lignos, Vatsal, and Scassellati (2006)

Study	Ability	Effect
Ishii, Nakano, and Nishida (2013)	tracking user's facial features	engagement awareness, likability, intelligence
Skantze, Hjalmarsson, and Oertel (2014)	tracking user's facial features	affects turn-taking
Zheng, Moon, Croft, and Meng (2015)	tracking user's facial features	likability, anthropomorphization
Baxter, Kennedy, Vollmer, de Greeff, and Belpaeme (2014)	tracking user's gaze	gaze to robot decrease over time
Lohan et al. (2011)	tracking user's gaze	appropriateness, responsive
Pitsch et al. (2009)	tracking user's gaze	sustained engagement
Szafir and Mutlu (2012)	monitoring and adjusting to user's engagement levels (EEG)	improved recall
Kidd and Breazeal (2008)	adapting greeting to the time of day	unclear effect
Klamer, Allouch, and Heylen (2010)	giving user a daily weather report	unclear effect
Sirkin, Fischer, Jensen, and Ju (2015)	displaying awareness to what happens outside the local interaction context	system is perceived as a serious interaction partner,increased trust and acceptability as a communication partner
Sirkin, Fischer, Jensen, and Ju (2016)	raising unexpected topics,displaying awareness to what happens outside the local interaction context	system is perceived as a serious interaction partner,increased trust and acceptability as a communication partner

Table 5.1: Summary of Literature Review for the Perceptual Basis

find that people interacting with a robot averting its gaze and gazing toward a user's face rate the robot as more enjoyable and more intentional than when it was static. In a study in which participants played a game with a robot, Rossi et al. (2015) find that the robot is rated as more natural and satisfactory when the robot was tracking participants' faces. While they did not employ active tracking, Fischer, Jensen, Suvei, and Bodenhagen (2016) test perceptive and interactional effects of robot gaze to participants during robot approach. They find that people find a large service robot to be more intelligent and more cooperative and that they are more at ease when the robot looks at them rather than looking at where it is going, while it approaches them. With regard to performance, a robot tracking participants has been found to increase participants' response times but to improve performance in easy tasks, but decrease performance in difficult tasks (Stanton & Stevens, 2014).

5.2.3 Displays of Awareness to the Perceptual Basis

Previous work concerning the perceptual basis has already been covered in Chapter 4 and is summarized in Table 6.1.

5.2.4 Summary

The review of the current literature indicates that a face-tracking mechanism contributes to the extent to which participants like the robot (Rossi et al., 2015; Wang et al., 2006) and the extent to which they find it intelligent. Results from Chapter 4 suggest that displays of awareness to the perceptual basis affect the degree to which participants find the robot social and interactive. Previous work on incrementality indicate that incremental feedback can affect the extent to which participants see a robot as polite and efficient (Baumann & Lindner, 2015; Skantze & Hjalmarsson, 2013), natural (Buschmeier et al., 2012), responsive, enjoyable, and attentive (de Kok et al., 2015; Tsai et al., 2018). The review suggests that either of these signals positively affect how people perceive robots, but does not enable predictions of the signals compare against each other. For example, both face-tracking and incrementality have been reported to affect likability.

5.3 Method

In the experiment, a robot (the EZ-Robot JD Humanoid) guides each participant through a series of physical exercises. Participants' perception of the robot are evaluated in a post-experiment questionnaire, while their behavior is evaluated by how much water they drink during the exercise, and the extent to which participants follow the robot's prompt to drink water.

5.3.1 Experimental Conditions and Manipulations

In this experiment, three contextual features are investigated in four experimental conditions. The three contextual features are awareness to *the perceptual basis*, *face-tracking*, and *incrementality*. One condition includes only the *face-tracking*, one condition includes only displays of awareness to the perceptual basis (referred to as *awareness*), one includes incremental feedback with face-tracking and awareness to the perceptual basis (referred to as *incrementality*), while the last condition includes none of these features (referred to as *none*). *Incrementality* was implemented during two exercises in which the robot provided online feedback on how participants were doing. In conditions without *incrementality*, the robot followed a set script and did not provide any feedback to participants. In the *face-tracking* condition the robot tracked a participant's face whenever it could find it. In conditions without *face-tracking*, the robot's head moved only as a result of other actions (e.g. nodding). Displays of awareness to *the perceptual basis* was implemented by having the robot comment on the weather (good or bad) and the robot commented on how much water was still in the respective participant's glass. In conditions without displays of awareness to *the perceptual basis*, the robot did not comment on the weather, and did not comment on how much water is in a participant's glas.

5.3.2 Participants

107 participants were recruited from the University of Southern Denmark, Campus Sønderborg. However, due to breakdowns in the connection between the Microsoft Kinect 2

and the computer controlling the robot, 23 interactions needed to be discarded. Of the 85 remaining participants, additional 5 had to be discarded because of other software breakdowns during the experiment. This leaves 80 interactions that are included in the final analyses. Mean age of participants is 27.2 (SD=10.1). Male participants are over-represented, so that 70% of the participants are men while the remaning 30% are women. However, these are balanced across the conditions.

5.3.3 Interaction Protocol

Participants were seated in front of the robot and the experiment began as the robot greeted them to the study. Next, the robot told participants its name, and asked for theirs and then asked them how they were feeling. Then the robot offered participants a glass of water by saying:

> "*Would you like a glass of water*"

In the *awareness* and *incremental* conditions the robot would also look towards and point towards a water jug. Next, the robot told participants some of the healthy benefits of drinking water and asked how much water the respective participants drank each day. In the *awareness* and *incremental* conditions the robot gaze participants a negative assessment if they said they drank less than 1.5 liters a day and a positive assessment if they said they drank more than that each day. In the other conditions they robot always responded with:

> "*You should try to drink more*"

In the *incremental* and *awareness* conditions, the robot commented on the weather, while looking out the window. Thus if the weather was good the robot would say:

> "*Talking about mood, this weather makes me really happy*"

and if the weather was bad the robot would say:

> "*Talking about mood, this weather makes me really depressed*"

Next, the robot talked about some of the healthy benefits of physical exercise, after which it asked participants to raise their arms three times. In the *incremental* condition the robot gave participants feedback on how they were doing they exercise. For example if participants' arms were not raised, or not raised high enough the robot would tell them which arm they needed to raise a bit more (see more on incremental feedback below). Next, the robot asked them to stand up and do five squats. Again, in the *incremental* condition the robot gave them incremental feedback on how they were doing (see more on this in the section on incremental feedback below).

After participants sat down again the robot asked:

> "*Aren't you thirsty?*"

In the *incremental* and *awareness* conditions, the robot also commented on how much water a participant still had in his or her glass. Next, the robot thanked participants for their time. In the *awareness* condition and after participants had completed all exercises the robot said:

> "*We were a great team today, you look fitter already*"

if participants completed only some of the exercises the robot said:

> "*Maybe we can improve on some of the tasks next time*"

and if participants did not complete any exercise the robot said:

> "*I'm still a bit sad that you didn't exercise with me*"

In all other conditions the robot merely said:

> "*We were a great team today*"

Finally, the robot offered participants to take a sheet with instructions for exercises they can do at home on their own time.. In the *awareness* and *incremental* conditions the robot also looked towards the sheet and pointed towards the sheet of paper.

5.3.4 Robot and Software

The robot used in the experiment is the EZ-Robot JD Humanoid, which was also used in Chapter 3 and Chapter 4. For the experiment a modified version of the Wizard-of-Oz Module introduced in Chapter 4 was used.

In addition to the internal VGA camera, the robot is also connected to a Microsoft Kinect (V2), which feeds back x,y,z coordinates of participants' hands and spine base. This information is used to assess participants' performance and compliance when carrying out exercises and to provide incremental contingent feedback to participants. This feature is also only activated in the incremental condition. This functionality is enabled by the EZ-Builder plugin *DepthSensor*, written by user ptp[2].

5.3.5 Speech Management

Speech was produced via a text-to-speech engine that relies on the Microsoft Speech API (SAPI), using the 'David' voice. Most of the speech was prescripted and synthesized on-the-fly as the robot reached the state for any given speech. These states were reached either temporally, set by a wizard, or as a result of a user action. Temporally reached states were ones in which the robot spoke a specific utterance after a predetermined amount of time. Wizard-controlled states are ones determined by a wizard, whose input was needed, for example to determine how much water was still in the glass or for timed events that could not be planned in advance. The robot was also able, by merit of the Kinect, to determine how well the participant was doing each exercise, and on this basis provide online feedback in the form of speech.

[2] https://www.ez-robot.com/EZ-Builder/Plugins/view/173

Incremental Feedback

The robot (in the incremental condition) gave feedback to how participants were performing the exercises. For the ‘raise arms’ exercise, the robot told the participant either to raise the right hand, the left, or both if the hands were not detected at the expected coordinates. A set of predetermined values on the y-axis were used to determine whether participants were raising their hands, whether they raised them only a bit, or whether they were not participating at all. The Kinect plugin sent data to the EZ-Builder program at a rate of between 20 to 30 frames per second, while the robot updated its state once every second. This worked through a simple loop in the EZ-Builder program:

```
if($handright > $inclimit AND $handleft < $inclimit)
  SpeakStop()
  say("raise the right arm a bit more")
  Sleep(1000)
  goto(startKinect)
endif
```

Source Code 5.1: Raise Hands Code

The values for limits were set during the piloting of the experiment and are based on the mean height participants could raise their hands in the pilot without straining themselves. Also the feedback given to participants in case the robot did not detect any movement was incremental. This was done by a loop that gave a new and more detailed type of feedback with every increment. After more than 10 seconds of inactivity, the robot gives up and abandons the activity:

```
if($handRight < $inactlimit AND $handLeft < $inactlimit)
sleep(3000)
  if($inactCount = 0)
    say("I can't see you moving")
    $inactCount++
    goto(startKinect)
  endif
  if($inactCount = 1)
    say("I still can't see you moving")
    $inactCount++
    goto(startKinect)
  endif
  if($inactCount = 2)
    say("Move both your hands all the way up, like me")
    $inactCount++
    goto(startKinect)
  endif
  if($inactCount = 3)
    say("nevermind then")
    ControlCommand("Auto Position", AutoPositionAction, "Relax")
    goto(end)
  endif
goto(startKinect)
endif
```

Source Code 5.2: No Activity Code

The incremental feedback for the squat exercise is similar to that of the previous exercise. However, in addition to the arms (see Figure 5.1) the robot also monitors the location (on the y-axis) of participants' spine base[3]. Limits for the spine base, as with the arms, indicate the extent to which participants are doing the squat as instructed, doing the squat, but not bending down at all, or not participating at all. The values for the limits for the spine are in contrast to the limit for the hands, adjusted to each participant. This is done by recording a baseline location of participants' spine base at the moment when they are standing up and are ready for the next instruction. In other words, as participants said they were ready, the Kinect recorded the location of their spine base on the y-axis and calculated how low they should go for the exercise to 'count' as a repetition, with the location of the spine base as reference:

```
:startKinect
saywait("are you ready")
$currentspeech = waitforspeech(5, "yes", "yeah", "sure")
if($currentspeech = "yes" OR $currentspeech = "yeah" OR $currentspeech = "sure")
  $spinebaseDef = $SpineBaseY
  $squadtarget = $spinebaseDef * 1.8
  say("okay, let's go!")
  goto(startSquat)
endif
else saywait("ok, say, begin, when you are ready")
$currentspeech = waitforspeech(30, "begin")
$quitSquat = $quitsquat+1
goto(startKinect)
```

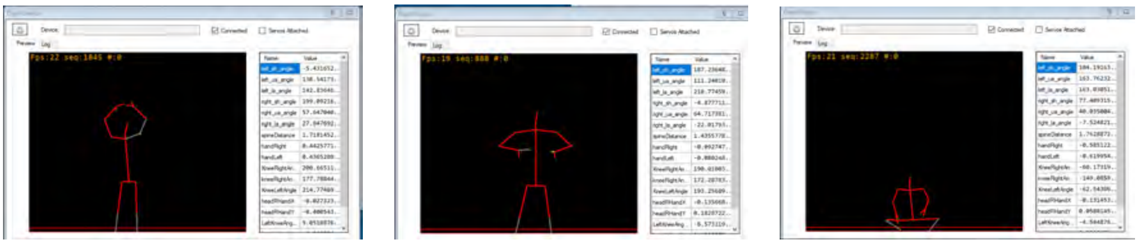
Source Code 5.3: Squat Target

The base value is multiplied by a factor of 1.8 to determine the target value for the spine base on the y-axis. The factor was found through experimentation with pilot participants and provides a target that requires some effort, without putting unnecessary strain on participants. While there are other methods that more accurately measure the extent to which participants are doing the exercise accurately, for example by measuring angles of the knee, pilot testing found the current method sufficient for the purpose. Figure 5.1 shows how joint locations are recognized by the Kinect.

5.3.6 Subjective Measures

Prior to the experiment, participants completed a questionnaire in order to get their demographic details, such as age, sex and their previous experience with robots. After the experiment, participants were asked to rate (on a 5-point scale) the extent to which the robot was *intelligent*, *authoritative*, *charismatic*, *strange*, *motivating*, *lifelike*, *trustworthy*,

[3]The spine base is the location in the skeletal tracker where participants' legs intersect with their spine (see Figure 5.1).



Raise arms | Squat baseline | Squat

Figure 5.1: Poses captured with the Microsoft Kinect

judgmental and *likable*. In addition, they were also asked about the extent to which they felt more *motivated* to do more exercise after completing the experiment, the robot *encouraged* them to complete the exercises, they felt *pressured by the robot*, they *felt watched by the robot*, they thought exercising with the robot was *fun*, and about the extent to which they tried to do the exercises as *accurately* as possible. These questions were designed to elicit the extent to which participants liked the robot, were motivated by the robot, and felt monitored by the robot.

5.3.7 Objective Measures

Two objective measures are analyzed: how much water participants drink during the experiment (in millimeter), and how participants respond to the robot's prompt to drink.

5.3.8 Analysis

A statistical analysis is performed on responses from the questionnaire using multiple linear regression, as well as for some of the objective measurement, taking into account effects of gender, previous experience with robots and the experimental condition. In addition, objective measures, such as how much water participants drink during the experiment, are subjected to statistical processing using multiple linear regression. Finally, how people respond to the robot's prompt to drink is analyzed using logistic regression.

5.4 Results

5.4.1 Subjective Measures

Analyses of the perceptive traits show significant effects of five of the nine traits under consideration. Figure 5.2 visualizes the statistically significant effects[4] of the predictors on the subjective measures. *Experience with robots* and *participant gender* affect the outcome variables only to a small extent and they do not interact with the experimental conditions. Specifically, participants with more experience with robots rate it as less *intelligent* than people who do not have much prior experience (B=0.37, SE=0.17, p=0.03). Men generally find the robot (regardless of condition) less *authoritative* than women do (B=0.47, SE=0.27, p=0.09).

[4]Full lines denote significant effects, while dashed line denote marginally significant effects

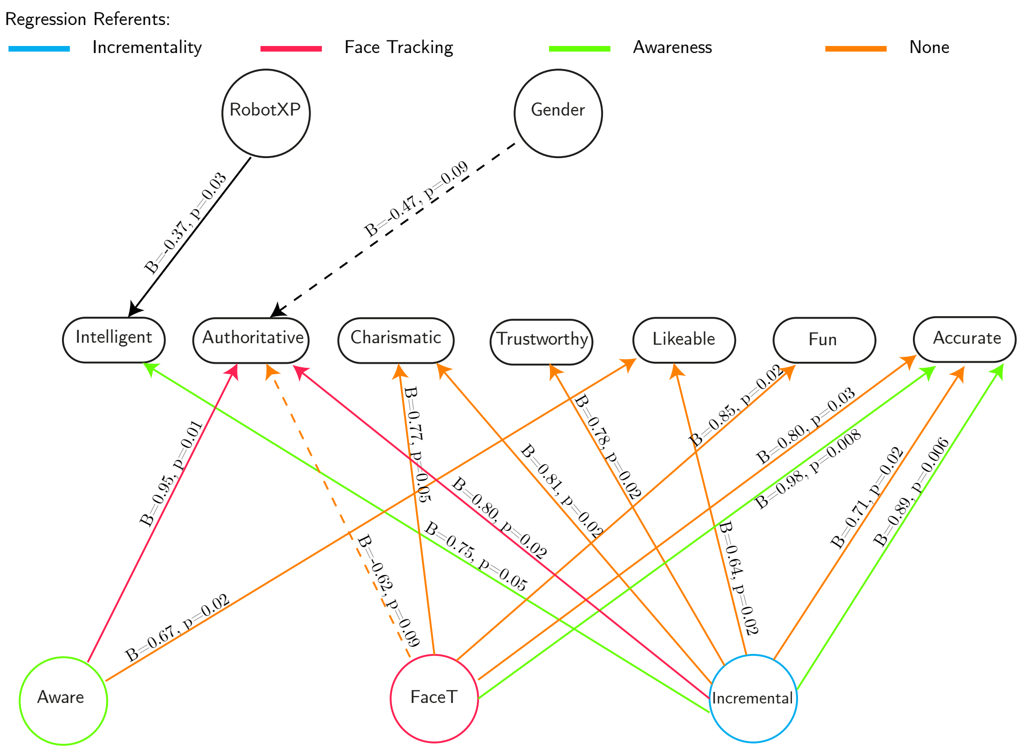


Figure 5.2: Regression Model

Face-tracking has a significant positive effect on *charisma* (B=0.77, SE=0.38, p=0.05) and on *authority* compared to *awareness* (B=-0.95, SE=0.37, p=0.01), *incrementality* (B=0.80, SE=0.45, p=0.02), and *none* (marginally significant, B=0.62, SE=0.26, p=0.09)[5]. *Awareness* positively influences the robot's *likability* compared to *none* (B=0.67, SE=0.28, p=0.02) (see Figure 5.3). *Incrementality* has a positive significant impact on *charisma* (B=0.81, SE=0.33, p=0.02), *trustworthiness* (B=0.78, SE=0.34, p=0.02), *intelligence* (B=0.75, SE=0.38, p=0.05), and *likability* (B=0.65, SE=0.27, p=0.02)[6].

Analyses of participants' ratings of the experiment show positive significant effects of *face-tracking* compared to no features on the degree to which participants thought exercising with the robot was *fun* (B=0.85, SE=0.35, p=0.02), and the extent to which they tried to do the exercises as *accurately* as possible (B=0.71, SE=0.35, p=0.02). The self-reported extent to which participants attempt to the exercises *correctly* is also significantly higher in the *face-tracking* condition compared to the *awareness* condition (B=0.98, SE=0.36, p=0.008), see Figure 5.4. Likewise, *incrementality* shows a positive significant effect of the extent to which participants tried to do the exercises as *accurately* as possible compared to *none* (B=071, SE=0.31 , p=0.02), and compared to *awareness* (B=0.89, se=0.31, p=0.006)

5.4.2 Objective Measures

Analysis shows that *awareness* (B=74.60, SE=33.60, p=0.03), as well as *incrementality* (B=83,19, SE=34.42, p=0.02) compared the *none* condition, positively and significantly

[5] Positive impact means a low score

[6] The full regression model is available in Appendix A

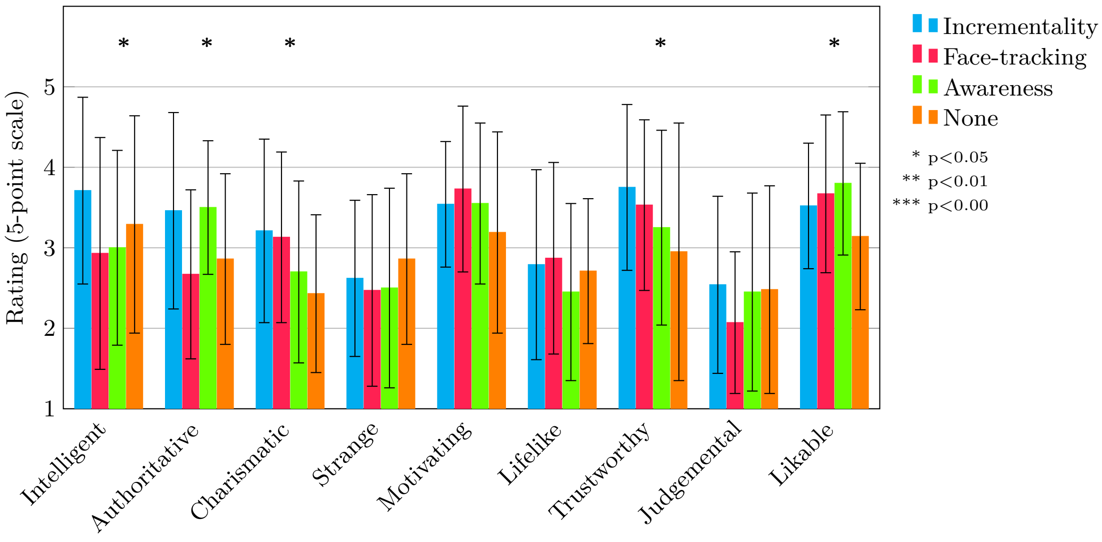


Figure 5.3: Perceptive Traits

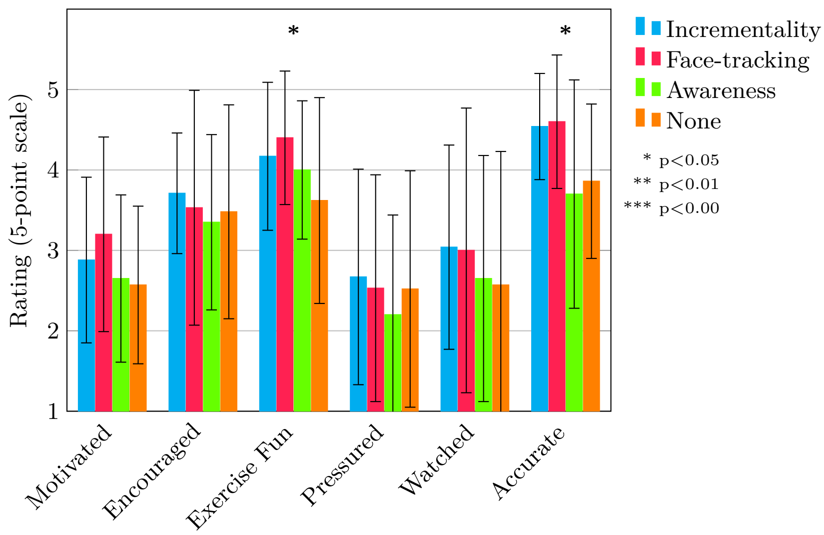


Figure 5.4: Experiment Evaluation

influence how much water participants drink (see Figure 5.5).

Likewise, *awareness*, as well as *incrementality*, positively and significantly influence the likelihood of participants drinking water when prompted to do so by the robot (see Figure 5.6). Initially, only participants interacting with the robot in the *incremental* condition are significantly more likely to drink when prompted to do so (B=1.46, SE=0.68, p=0.03). The second time participants are prompted to drink, both participants interacting with the robot in the *incremental* condition (B=2.40, SE=0.79, p=0.002) and to an even higher degree participants who interact with the robot in the *awareness* condition (B=30.2, SE=0.87, p=0.0005) are significantly more likely to drink compared to the *none* condition. Likewise, participants in the *awareness* condition also are significantly more likely to drink than participants in the *face-tracking* condition (B=1.83, SE=0.82, p=0.02)

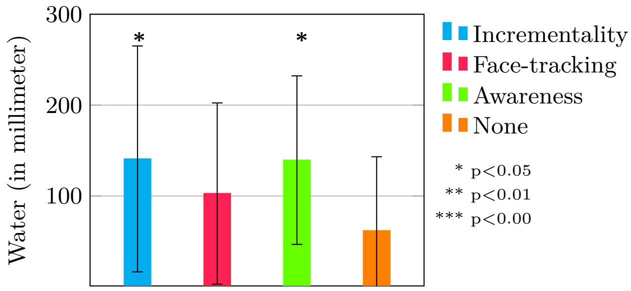


Figure 5.5: Water Consumption

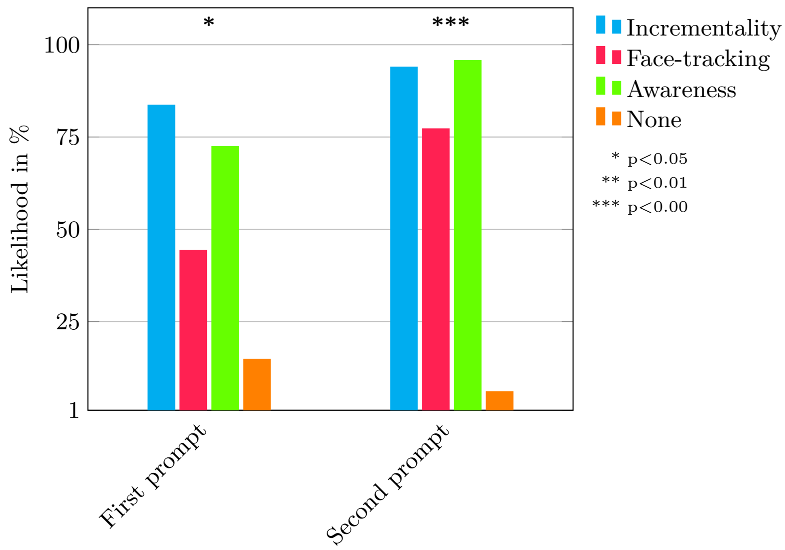
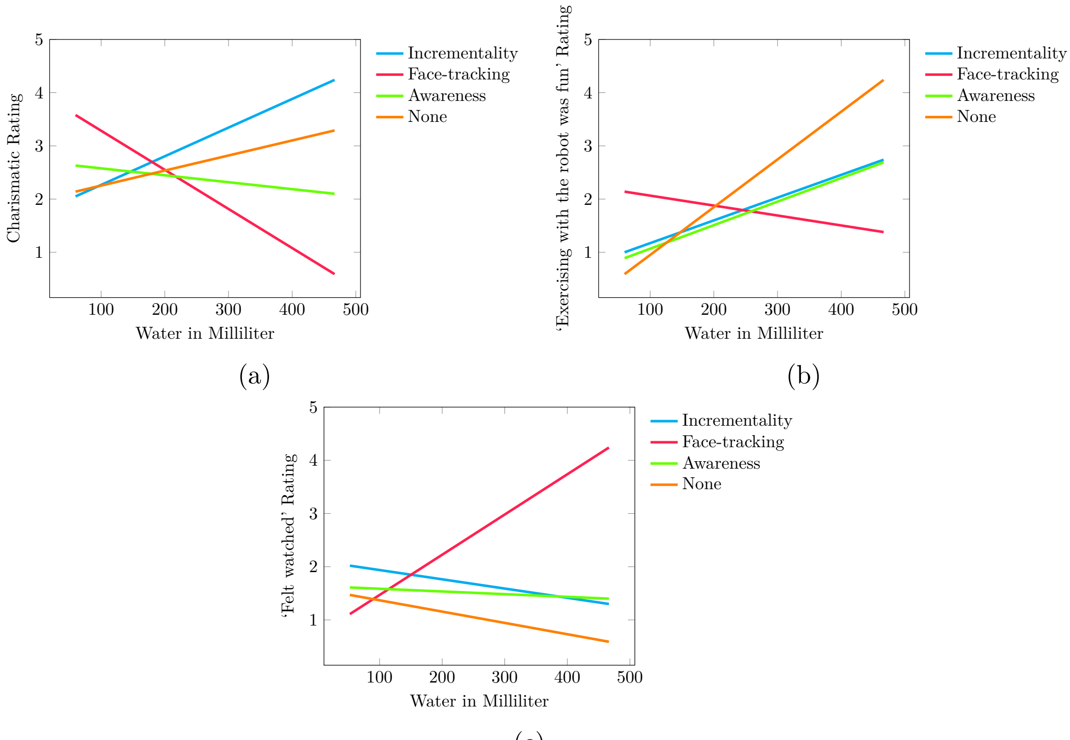


Figure 5.6: Prompt to Drink

5.4.3 Interactions Between Measures

Interaction analyses between subjective measures (i.e. questionnaire responses) and how much water participants drink during the experiment reveal three interesting results. First, analyses reveal a possible trade-off effect between the *face-tracking* feature and participants' perceptions of the robot and experiment. Specifically, the charismatic rating and water consumption are negatively correlated in the *face-tracking* condition (B=–0.0095, SE=0.0042, p=0.02). That is, the more water participants drink, the less charismatic they perceive the robot to be (see Figure 5.7(a)). A similar relationship is found between the degree to which participants found it *fun* to do the exercises with the robot and how much they drank in the *face-tracking* condition (B=-0.0091, SE=0.0038, p=0.02). Again, the more water participants drink, the less fun they thought it was doing the exercises with the robot (see Figure 5.7(b)). Also the degree to which participant felt watched and how much water they drank during the experiment interact significantly with the *face-tracking* condition (B=0.013, SE=0.0062, p=0.03). That is, the more participants felt watched in the *face-tracking* condition, the more water they drank (see Figure 5.7(c)). Thus, these results seem to suggest a trade-off between behavioral compliance and perceptive traits.

Second, analyses reveal that whether or not participants find the robot motivating becomes less important for their water consumption when the robot uses either of the three contextual

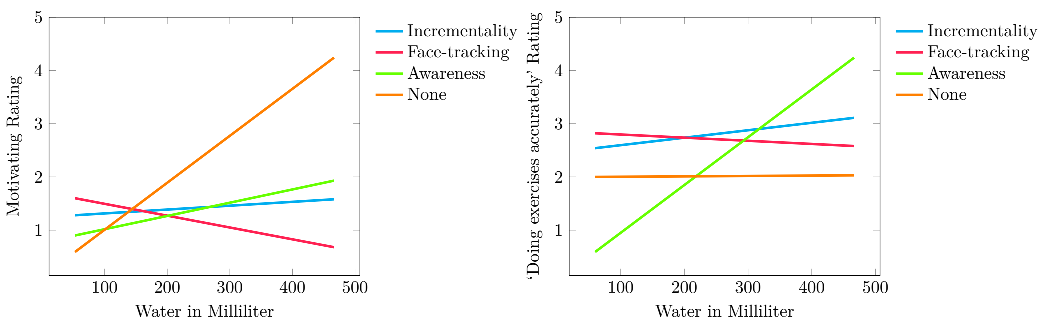


(a) (b) (c)

Figure 5.7: Interactions I

features (see Figure 5.8(a)). Whether or not participants find the robot motivating seems to have only little effect on their water consumption in either of the three conditions. However, whether or not participants find the robot motivating in the *none* condition influences participants' water consumption significantly. Specifically, the *incrementality* (B=-0.0085, SE=0.0033, p=0.01) and the *face-tracking* conditions (B=-0.011, SE=0.0038, p=0.005) differ significantly from *none* condition, whereas the difference is only marginally significant (B=-0.0068, SE=0.0037, p=0.07) between the *awareness* condition and the *none* condition.

Third, participants' water consumption is positively correlated with the degree to which they say they attempted to do the exercises as accurately as possible (see Figure 5.8(b)), but only for participants' in the *awareness* condition (B=0.0077, SE=0.0038, p=0.05), while ratings in the other conditions seems unrelated to participants' water consumption.



(a) (b)

Figure 5.8: Interactions II

5.4.4 Results Overview

All significant differences reported above are summarized in Table 5.2. Since the incremental condition also includes the manipulations from the face tracking and awareness conditions, it can be assumed that overlapping effects between the incremental condition and the other two conditions are due to to either face tracking or awareness, while any non-overlapping effects can be ascribed to incrementality. For example, Figure 5.2 shows that both the face tracking and the incremental conditions affect charisma. However, since the face tracking feature is also present in the incremental condition, it can be safely assumed that the effect is a function of the face tracking mechanism.

	Face Tracking	Awareness	Incrementality
Intelligence			X
Authority		X	
Charisma	X		
Trustworthinesss			X
Likability		X	
Fun	X		
Encourage Accuracy	X		
Water Consumption		X	
Prompt Efficiency		X	

Table 5.2: Summary of Significant Results

5.5 Discussion

5.5.1 Effects of Face Tracking

The study shows that a robot's ability to track participants with head movements makes people perceive it as more charismatic, and people find the robot more fun to interact with. In addition, participants report that they attempt to a greater extent to do the exercises accurately. The results for charisma and fun confirm previous works, which have shown that face tracking contributes to enjoyment (Wang et al., 2006) and satisfaction (Rossi et al., 2015). These results also resonate well with the findings from the Chapter 3, which suggest that participants rate the robot as more friendly in the contingent gaze condition. While face-tracking and contingent gaze are indeed two different kinds of behaviors, they look very similar from a participant's point of view.

The finding that face tracking also encourages participants to do exercises more precisely is new, but compatible with previous work. For example, Stanton and Stevens (2014) found that a when participants interacted with a robot that tracked them, participants were found to increase performance in easy tasks, but to decrease performance in more difficult tasks. The authors argue that a robot's gaze exerts pressure on participants to perform, and that this leads to increased performance for tasks that are easy to do. The incremental behavior, which provides feedback on participants' performance, does not encourage accuracy more than face tracking. In fact, face tracking slightly outperforms incrementality, although this difference is not statistically significant.

While no main effects for objective measures are found, there are several significant interaction effects between participants' water consumption and their perception of the robot. Specifically, analyses revealed that participants who drank more water also felt more watched, had less fun, and rated the robot as less charismatic. Face tracking is therefore, as with other manipulations of gaze, not a simple matter. Because of the multifunctional nature of gaze, which was documented in Chapter 3, it is difficult to ascertain exactly which gaze behaviors are appropriate, at which time during an interaction.

5.5.2 Effects of Displays of Awareness of the Perceptual Basis

Awareness is the only condition that yields significant main effects for both subjective and objective measures. Specifically, participants found the robot more likable compared the *none* condition, and also more authoritative compared to the *face-tracking* condition. The robot is able to adjust to states outside the local interaction context (the weather). It is thus able to engage in small talk, which is likely the reason why participants rate it as more likable. This result also resonates well with previous work that has found robots that are able to remember and produces utterances based on participants' language production, to be rated as more appropriate (Zhang et al., 2018), likable (Leite et al., 2017) as well as social and interactive Chapter 4.

However, participants also found the robot to be more authoritative in the *awareness* condition. This is likely caused by the reference the robot makes to how much water is still in participants' glasses. It is interesting to note that while participants found the robot more authoritative in the *awareness* condition, they do not regard it as judgmental. This suggest that the manipulation is powerful enough for people to notice, but not so powerful that it impacts perceptions negatively. Analyses of the objective measures also support this argument. Participants comply more positively with the robot's prompt to drink more water in the *awareness* condition, and also drink more water during the experiment. Verbal references to the perceptual basis of are very effective means to nudge people to particular behaviors.

5.5.3 Effects of Incrementality

The current study shows that incremental feedback makes participants perceive the robot as more intelligent (in comparison to the *awareness* condition) and more trustworthy (in comparison to the *none* condition). Previous work that has used incremental feedback for similar exercises has also found that incremental feedback positively influences participants' perception of the robot's intelligence. Previous work has found incremental feedback to positively influence politeness, efficiency (Baumann & Lindner, 2015; Skantze & Hjalmarsson, 2013), responsivity, enjoyability, and attentiveness (Tsai et al., 2018). However, no studies have yet reported effects on trustworthiness.

5.5.4 Conclusion

The current study shows that each of the displays of awareness of contextual information contributes in different ways to participants' perceptions of the robot and to the interactions themselves. There is very little overlap between conditions, which serves to show that each of the displays contribute differently. That is to say, the kind of contextual information a robot displays an awareness towards has a large impact on how it is perceived and responded to.

The awareness condition was the only one that was shown to directly affect participants' conduct, but interactions between subjective measures and water consumption showed that face tracking can lead to less desirable outcomes, such as the feelings of being watched, and a perceived low level of charisma. In the current analysis, it is however not possible to ascertain exactly why this is the case.

6. Study 4: Incrementality

6.1 Introduction

This chapter investigates perceptual and interactional effects of incremental feedback. In interaction, incremental speech can display an awareness to the conduct of communication partners or to perceptual experiences. As such, incremental feedback forms part of of what one's communication partner take the common ground to be. In Chapter 5 the robot displayed an awareness to participants' hand locations when participants failed to lift their hands high enough. This display, or grounding, was was continuously updated on a moment-by-moment basis, i.e. incremental, and reflected what participants were doing with their hands. The robot displayed to participants in real-time that it considered their bodily movements as part of the common ground that existed between them.

From an engineering perspective, there are (at least) two arguments for why robots should produce speech incrementally. One is technical, and sees incremental speech as a way to improve the efficiency of system, by for example initiating utterances before they have been fully realized by the system. Since production happens before processing is complete, such a mechanism can reduce the time it takes for a system to produce a response (Pouget, Hueber, Bailly, & Baumann, 2015; Skantze & Hjalmarsson, 2010; 2013; Tsai et al., 2018). The other argument rests in the fact that the way people understand and produce speech is incremental (Goodwin, 1979), and that robots that engage people in social interaction should do so too, (Baumann & Lindner, 2015; Buschmeier et al., 2012; Scheutz, Cantrell, & Schermerhorn, 2011; Yu, Bohus, & Horvitz, 2015).

Systems, and in particular robots that are able to produce incremental speech, are evaluated on the basis of either performance or perception. Performance generally means that a system is able to respond faster, or that it enables people to do their tasks faster or better than without incrementality. Perception is generally investigated through post-experiment questionnaires. However, the relation between perception and performance is rarely ever investigated.

The current study aims to investigate effects of incremental verbal feedback on perception, performance and the interaction between the two. Specifically, I investigate how a robot's displays of awareness to participants' conduct in an object-searching task affect participants' efficiency when solving the task, and how they perceive the robot. In addition, I also investigate how perception and performance interact. The kind of incremental feedback implemented is similar to what is implemented by de Kok et al. (2015) and to what was implemented in Chapter 5.

Chapter 5 investigated, among other things, how incremental feedback changes how people perceive a robot. The study indicated that incremental feedback positively affected the extent to which participants perceived the robot to be intelligent and trustworthy (see Table 5.2). The current chapter follows up on these results on a different robot. Two factors inform this decision. First, I want to see the extent to which results from the prior study are stable across robot platforms. Second, as the previous chapter only investigated perceptual effects of incremental feedback, I want to also investigate how incremental feedback affects interaction between participants and a robot.

6.2 Literature Review

A literature review of incrementality has already been carried out in Chapter 5. The main findings are summarized in Table 6.1.

Study	Focus	Contribution of incrementality
Baumann and Lindner (2015)	Perception	incremental speech systems are perceived as more polite and efficient (polite and natural)
de Kok et al. (2015)	Perception	incremental speech systems are perceived as more intelligent, helpful, responsive, clear and tiring
Skantze and Hjalmarsson (2013)	Perception	incremental speech systems are perceived as more polite and efficient
Tsai, Baumann, Pecune, and Casell (2018)	Perception	incremental speech systems are perceived as more responsive, enjoyable, and attentive
Buschmeier, Baumann, Dosch, Kopp, and Schlangen (2012)	Perception	incremental speech systems are perceived as more natural
Chromik, Carlmeyer, and Wrede (2017)	Perception	simulated robot using natural speech is perceived as less natural
Carlmeyer, Schlangen, and Wrede (2016a)	Perception	simulated robot using incremental speech is perceived as less likable
Kennington et al. (2014)	Performance	incrementality (incremental feedback) increases performance (better driving)
Skantze and Hjalmarsson (2010)	Performance	incrementality (incremental speech) increases performance (response time)
Skantze and Hjalmarsson (2013)	Performance	incrementality (incremental speech) increases performance (response time)
Ghigi, Eskenazi, Torres, and Lee (2014)	Performance	incrementality increases dialogue length and increases success rate (information retrieval)
Chromik, Carlmeyer, and Wrede (2017)	Performance	incrementality improves performance (object retrieval)

Table 6.1: Literature Review

6.3 Method

6.3.1 Experimental Conditions

The experiment is carried out in a between subject experimental design with two experimental conditions. In one condition, the robot is able to modify its speech incrementally based on participants' non-verbal conduct. As participants are looking for certain hidden items, the robot can direct their search by producing utterances like "more to the right" and "yes a little more". Thus, the incrementality speech processing implemented for this experiment is in effect very similar to the one implemented for the experiment in Chapter 5. This condition is referred to as the incremental condition. The other condition, referred to as the non-incremental condition, features no incremental speech. Instead, the robot is only able to repeat its previous utterance in cases where participants are not finding what they are looking for.

6.3.2 Subjective Measures

Participants were presented a questionnaire before and after the experiment. Demographic information and previous experiences with robot were elicited in the questionnaire given prior to participating, while participants' ratings of the robot were elicited in the second questionnaire. The post experiment questionnaire is not identical to the one given in Chapter 5, but consists of the RoSAS scale (Carpinella, Wyman, Perez, & Stroessner, 2017) as well as some additional questions. However, items in the RoSAS scale are comparable to the items given in the Chapter 5. The motivation for this change was to have a measuring instrument, which measures social perception, which it claims to do (Carpinella et al., 2017). Other instruments, such as the Godspeed series (Bartneck, Kulić, Croft, & Zoghbi, 2009), are more concerned with measuring anthropomorphism, which is not directly related to the research questions I pursue.

The scale consists of the three indices warmth, competence, and discomfort. Each index contains a collection of adjectives with which the participants are asked to rate the robot. All items are rated on a 5-point likert scale where 1 is 'not at all' and 5 is 'very much'. Warmth includes the adjectives *happy*, *feeling*, *sociable*, *compassionate*, and *emotional*. Competence includes the adjectives *capable*, *responsive*, *interactive*, *reliable*, *competent*, and *knowledgeable*. The discomfort scale includes the adjectives *scary*, *strange*, *awful*, *awkward*, *dangerous*, and *aggressive*. Cronbach's Alpha was calculated for each of the indexes. The competence and warmth index both score an alpha of 0.82, while the discomfort index scores an alpha of 0.66. While index scores for warmth and competence suggest that these are reliable, the index score for discomfort suggests that it does not reach the same level of reliability. In addition to this scale another five adjectives were added. These were *boring*, *credible*, *engaging*, *likable*, and *enthusiastic*. Finally, three questions were added to work as manipulation checks, to be evaluated on a five-point scale as well. These questions were:

- to what extent do you think the robot took you into account?
- to what extent do you think the robot responded to your actions?
- to what extent do you think the robot perceived you?

6.3.3 Objective Measures

The objective effect of incremental speech is evaluated by measuring the time it takes participants to find the two objects that are concealed from view. Time is measured from when the robot issues the instruction and until participants takes hold of the object.

6.3.4 Hypotheses

The literature review indicates that incrementality positively influences factors related to competence, such as intelligence (de Kok et al., 2015), and efficiency (Baumann & Lindner, 2015; Skantze & Hjalmarsson, 2013). Likewise, Chapter 5 indicates that incrementality positively influences intelligence and trustworthiness. I therefore hypothesize that:

H_1: Participants rate the robot in the incremental condition as significantly more competent and more credible

Previous works on affective factors are inconclusive (see the literature review), so results for boring, engaging, likable, and enthusiastic are not hypothesized. Previous work indicates that incrementality has a positive effect on efficiency (Kennington et al., 2014; Skantze & Hjalmarsson, 2010; 2013), thus I hypothesize that:

H_2: Participants find objects faster in the incremental condition than in the non-incremental condition.

Previous work on incrementality focus on either performative or perceptive metrics, but virtually none explore the relationship between the two. However, a robot that can help people perform a task more efficient is likely to be perceived as more competent. Thus, the third hypothesis is that:

H_3: The faster participants find the objects in question, the more competent they will find the robot.

6.3.5 Interaction Protocol

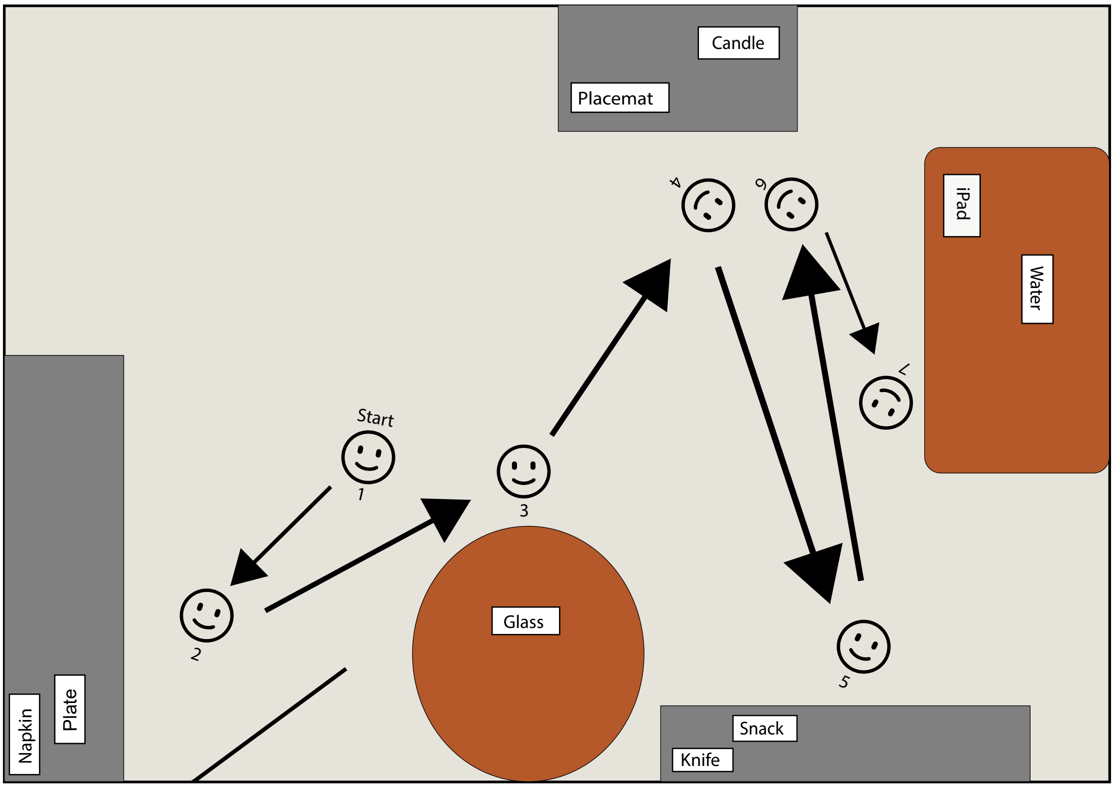


Figure 6.1: Experiment Map

Figure 6.1 describes the setup of the experiment in the lab. Robot and participant start (1) where the participant is greeted by the robot. Next, the robot moves off to (2), where the robot instructs the participant to pick up a plate and a napkin. In the incremental condition, the robot uses incremental speech to direct the participant to the napkin, which is concealed from view. Then, the robot moves off to (3) and instructs the participant to pick up a glass, continues to (4) where the participant is asked to pick up one of two placemats. Here the robot explicitly displays its situation awareness by commenting on the placemat the participant picks up. For example if the participant picks up the green placemat the robot says ‘Ah the green one, that’s my favorite too!’. If no items have yet been placed on the robot, it offers to carry them for the participant. Next, the robot moves off to (5) and instructs the participants to pick up a snack. Here, participants can choose between a cookie or a fruit. Again, the robot displays situation awareness by commenting on their choice (however, judging their choice). Then the robot moves on to (6) where participants are asked to pick up a candle. The candle is hidden away in a drawer, so for participants in the incremental condition the robot directs them using incremental speech. Finally, the robot goes to (7) and directs people to set the table, and have a seat, enjoy their snack and fill out the post experiment questionnaire that is prepared on a tablet (see Figure 6.2).

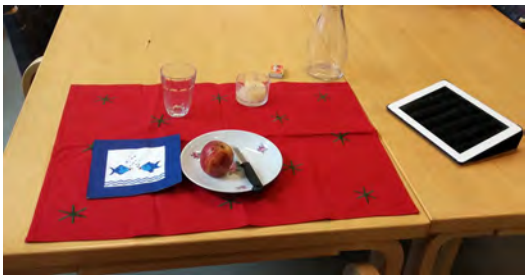


Figure 6.2: Set Table

6.3.6 Robot and Software

Figure 6.3: Robot

The robot used for this experiment is a Turtlebot 2 on a Yujin Kobuki mobile base. The robot is equipped with an Orbbec Astra 3D camera and is controlled by an Intel NUC running Canonical Ubuntu 16.04 LTS and ROS Kinetic. The robot moves autonomously from point to point. However, target locations are set by a remote wizard using RViz (see Figure 6.4). The autonomous navigation is enabled by SLAM map building (Pajaziti, 2014). The robot's speech is presynthesized using IVONA TTS (voice 'George'). A remote wizard controls, via a collection of shell scripts, when the robots produces its speech, and to a limited extent, what it says (because some actions are predifined, see below). Cameras are placed around the room on walls and on ceilings, and are live streamed to a PC in an adjacent room. The robot is a low-fidelity prototype, coated in styrofoam and equipped with a pair of eyes made from bottle caps (see Figure 6.3).

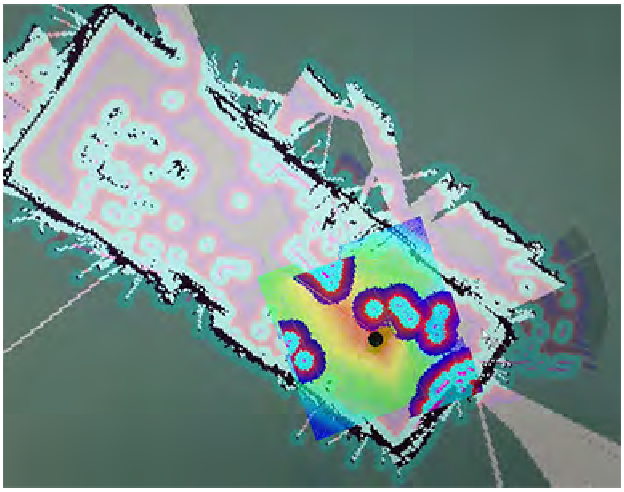


Figure 6.4: The Experiment Space Map

Speech Management

The robot's speech is presynthesized using IVONA TTS (voice 'George'). All of the robot's verbal actions are controlled via a series of shell scripts, thereby limiting the options available to the wizard at any given time during the experiment. This was to decrease the cognitive load of the wizard who already had to monitor several aspects of the participants' behavior and point the robot in the right direction. The script runs in the command line and the wizard select the next utterance by clicking the appropriate numerical key. This is demonstrated in Listing 6.1 below.

```bash
#!/bin/bash
control="0"

###Plate task
echo "1: Continue"
while true; do
read -rsn1 input
if [ "$input" = "1" ]; then
echo " "
echo "Our first task is to get a plate from this shelf over there"
echo "CONTROL: turn robot around and move toward shelf"
echo " "
play LivingLabAudio/plateNew.mp3
break
fi
done
```

Source Code 6.1: Sample Shell Script

Here, the '1.Continue' is printed to the screen indicating that the only option available to the wizard is '1', which will progress the interaction. Thus, in the example the wizard is not able to decide what the robot is going to say, merely when it is going to do so. The script ignores all other inputs than the numerical key press '1'. There are however also situations in which the robot needs to adjust to the participant's behavior. Such situations are resolved by giving the wizard a small list of possible actions to perform:

```
while [ "$control" = "0" ]
do
echo "1: User puts glass on robot"
echo "2: User keeps the glass"
echo "3: User does not pick up glass"
echo " "
while true; do
read -rsn1 input
if [ "$input" = "1" ] && [ "$place" = "0" ]; then
echo " "
echo "You are welcome to put everything on my tray."
let "place++"
echo " "
play LivingLabAudio/napkins6.mp3
control="1"
break
elif [ "$input" = "1" ] && [ "$place" > "0" ]; then
echo " "
echo "Great"
echo " "
play LivingLabAudio/great.mp3
control="1"
break
elif [ "$input" = "2" ]; then
echo " "
control="1"
break
elif [ "$input" = "3" ]; then
echo "please pick up a glass"
play LivingLabAudio/glass5.mp3
break
fi
done
done
```

Source Code 6.2: Sample Shell Script

Note that the wizard is not given options for what to do, but rather a list of possibilities that the participant could be doing. Thus, the wizard does not need to evaluate the 'right' course of action, but merely respond to what he or she is observing. The possible of participant behaviors are derived from pilot studies. The incremental speech is implemented in a similar fashion. That is, the script will keep looping the options available to the wizard until the participant reaches the target object. This creates the illusion that the robot incrementally adjusts its own verbal output to the participant's nonverbal conduct for example, by saying 'higher' when a participant needs to look higher up on a shelf. The scripts are run remotely from an adjacent room to the Turtlebot NUC via SSH.

6.3.7 Analysis

Subjective measures are evaluated using multiple linear regression, taking into account effects of sex, previous experience with robots and the experimental condition. The objective measure is evaluated by means of a general linear mixed model. The outcome variable is the time in seconds it took to find the object the robot asks for. Predictor variables are the experimental condition, sex, and previous experience with robots. Since each participant was asked to find two different hidden objects in which the robot produced either incremental or non-incremental feedback, participants' ID was entered into the regression as a random factor. The analysis is conducted using R 3.5.0 (R Core Team, 2017). Graphs presented are based on the means and standard deviations of each variable between the two experimental conditions. Regression coefficients (B), standard error (SE) as well as the probability value (p) are reported for the regressions under discussion.

6.3.8 Participants

51 students and staff from the University of Southern Denmark, Campus Sønderborg agreed to participate in the study. The staff includes both members of faculty, students' parents, as well as in-house personell. Students range from second semester bachelor level to ph.d.-students. Mean age of the participants is 28.2 (SD=11). Men are overrepresented in the study, so that only 31% are women, while the rest are men. However, these are balanced between the two experimental conditions. Participants with previous experience with robots are likewise equally distributed between the two experimental conditions. None of the participants have previously interacted with this robot before. Most participants are native speakers of either Danish or German, but there are also participants natively speaking Albanian, Bulgarian, Catalan, Chinese, Farsi, Frisian-Dutch, Greek, Hungarian, Icelandic, Japanese, Latviam, Malayam, Marathi, Polish, Russian, Spanish and Ukranian.

6.4 Results

6.4.1 Manipulation Check

Analysis of the three questions that make out the manipulation check reveals that participants generally rate the non-incremental robot more positively. However, neither 'responded to your actions' (B=0.16, se=0.29, p=0.60), 'into account' (B=0.38,se=0.28 p=0.18) nor 'perceived you' (B=0.22, se=0.27, p=0.40) are statistically significant. Previous experience with robots and participant gender had no effect of results.

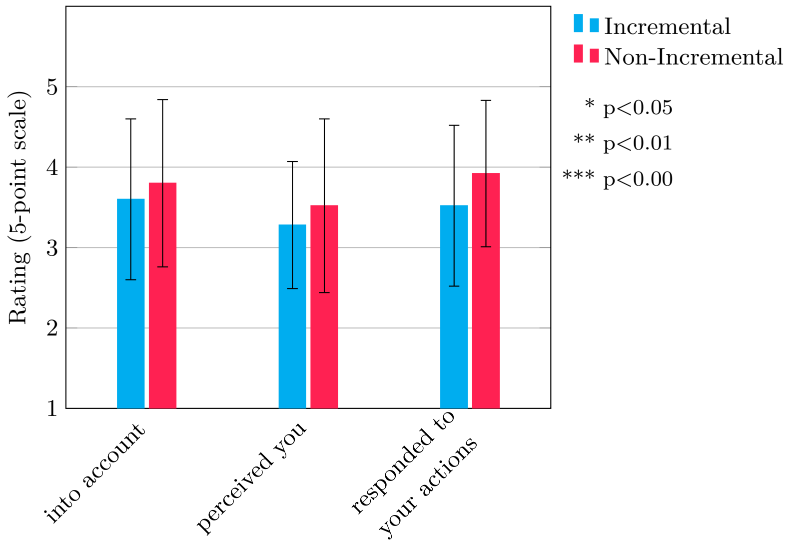


Figure 6.5: Manipulation Check

6.4.2 Questionnaire

The analysis of the RoSAS scale reveal no significant differences between the two experimental conditions. Thus, H_1 is rejected. However, for *competence* a significant difference for gender is found (B=-0.47, se=0.23, p=0.04). No effects of previous experience with robots are found.

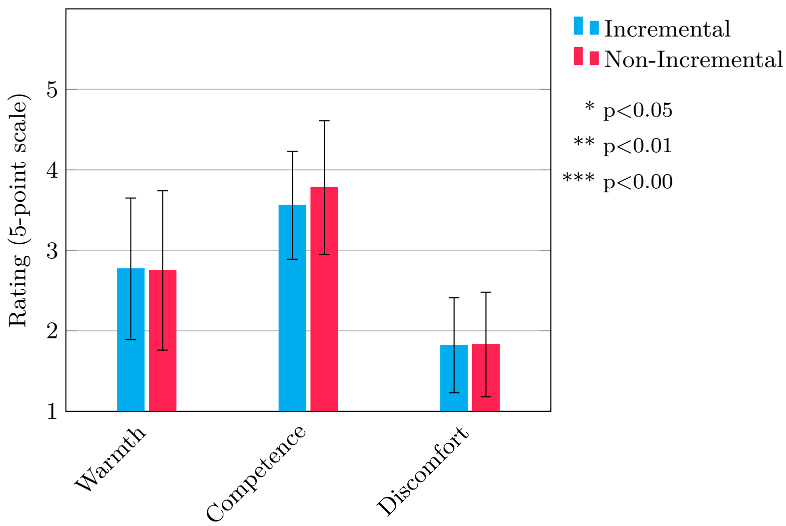


Figure 6.6: RoSAS Scale

Concerning the remaining questionnaire items, participants find the robot slightly more *engaged* and *enthusiastic* and less *boring* in the incremental condition while they find the robot more *likable* and more *credible* in the non-incremental condition (see Figure 6.7. However, only the differences between *credible* is statistically significant (B=-0.60, se=0.26 p=0.02).

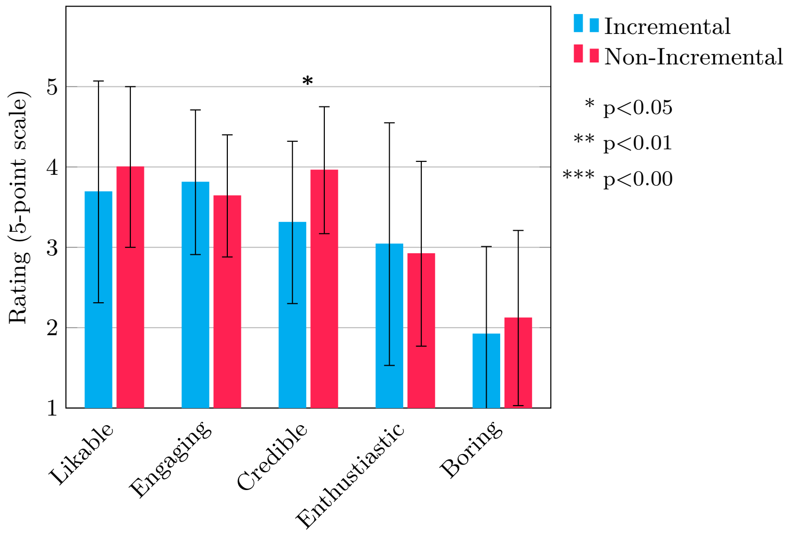


Figure 6.7: Additional Questions

In order to investigate effects of participant gender more closely the model is rerun with an added interaction term between the experimental condition and participant gender. This analysis reveals a significant interaction for the factor *warmth* (B=-1.45, se=0.54, p=0.01, see Figure 6.8). Specifically, while men rate the robot more positive in the incremental condition, the opposite is true for women. Women rate the robot more positive in the non-incremental condition than in the incremental condition. The interactions for *competence* (B=-0.48, se=0.46, p=0.30) and for *discomfort* (B=-0.35, se=0,39, p=0.38) are not statistically significant.

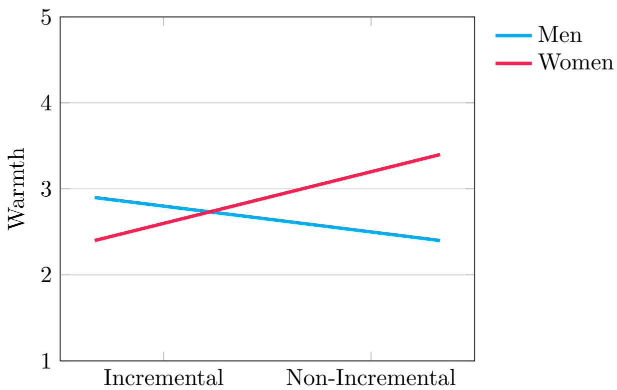


Figure 6.8: Warmth Interaction with Gender

6.4.3 Effectiveness

Analysis of the objective measure, namely how much time it takes participants to find the concealed items, reveals a statistical significant difference between the two experimental conditions. Specifically, participants in the incremental condition complete the task around six seconds faster than participants in the non-incremental condition (B=5.87, SE=2.43, p=0.02):

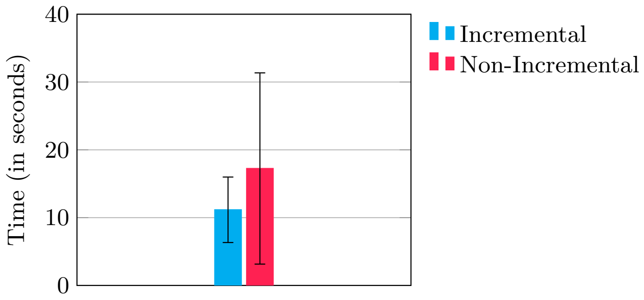
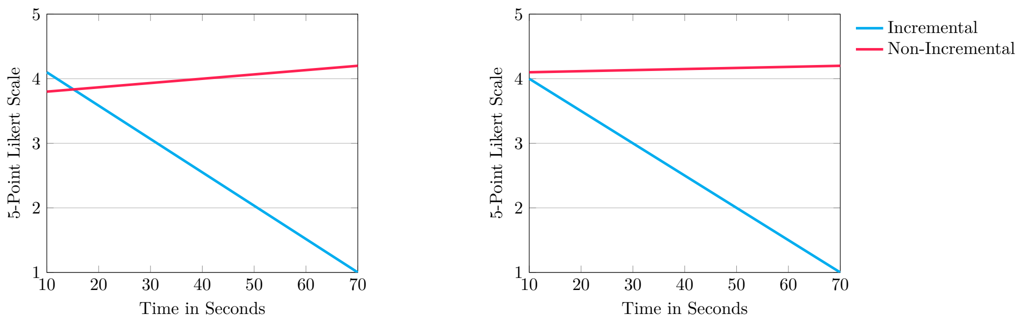


Figure 6.9: Effectiveness

6.4.4 Interactions Between Performative and Perceptive Metrics

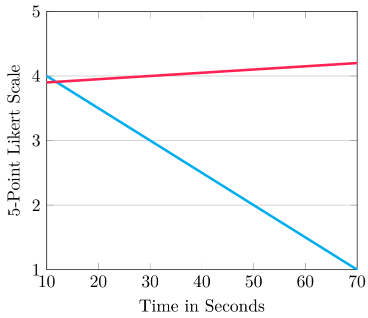
In order to explore relations between the performative metric and the perceptive metrics, a multiple regression is modeled with with the objective measurement (the time it took to find either of the two objects), experience with robots, participant gender, and the experimental condition as predictors, and perceptive measurement as outcome variable. In addition, an interaction term between the experimental condition and the objective metric is added. The results for the manipulation checks show a significant negative main effect of the time it took to find objects and the extent to which participants thought the robot responded to their actions (B=-0.07, SE=0.03, p=0.03), and the extent to which participants thought the robot took them into account (B=-0.07, SE=0.03, p=0.04). In other words, the longer it took for participants to find the right objects, the less they thought the robot took them into account and responded to their actions. Both the extent to which participants thought the robot responded to their actions (B=0.07, SE=0.03, p=0.03) and the extent to which participants thought the robot took them into account (B=0.08, SE=0.04, p=0.03) interact with the experimental condition (see Figure 6.10).



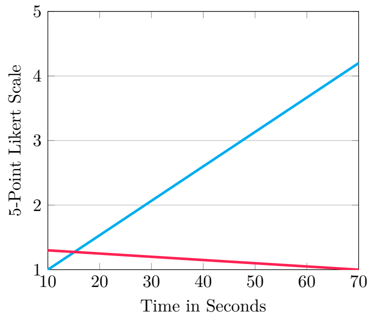
(a) Into Account (b) Respond to Actions

Figure 6.10: Interactions Between Performance and Perception I

For the indices in the RoSAS questionnaire, the time it took to find the objects had no effect on warmth (B=-0.04, SE=0.03, p=0.15), and did also not interaction with the experimental condition. A negative marginal effect is found for competence (B=-0.04, SE=0.03, p=0.09), and a positive marginal effect is found for discomfort (B=0.04, SE=0.02, p=0.09). The effect for competence interacts with the experimental condition (B=-0.05, SE=0.03, p=0.08), although the result does not reach statistical significance. Likewise, the effect for discomfort also interacts with the experimental condition (B=-0.04, SE=0.02, p=0.05). This result does reach statistical significance. Interaction effects for competence and discomfort are plotted in Figure 6.11.



(a) Competence



(b) Discomfort

Figure 6.11: Interactions Between Performance and Perception II

While analyses do show some relations between performative and perceptive metrics, the relation between competence and the time it took to find objects did not reach statistical significance. As such H_3 is rejected.

6.5 Discussion

This study aimed to investigate perceptive and performative effects of incremental feedback, and to investigate the relation between the two. The study confirmed previous work on how incrementality contributes to performance. However, for the subjective measures almost no differences could be found. Lastly, the interaction analyses between performance and perception reveal a relationship between the experimental conditions and the measurements, in particular in the degree to which participants thought the robot took their actions into account and the extent to which they thought the robot responded to their actions.

6.5.1 Effects on Perception

The study found almost no differences between the conditions. The one variable that was found to differ significantly between conditions showed that participants found the robot in the non-incremental condition more credible than participants in the incremental condition. This results stands in contrast to the study conducted in Chapter 5 in which participants rated the robot in the incremental condition as more trustworthy than participants in the non-incremental condition. Previous works have shown a relation between incremental speech and competence (Baumann & Lindner, 2015; de Kok et al., 2015; Skantze & Hjalmarsson, 2013) and incremental speech and affective factors (Tsai et al., 2018).In addition, many participants who interacted with the incremental robot expressed a fascination with the fact that the robot was able to guide them quite well to the objects they were looking for. It is therefore surprising to observe so few differences between conditions. In both conditions the robot was rated relatively positively. Therefore an explanation for the lack of differences can perhaps be found in that fact that the robot, in either condition, performed reasonably well, and acted contingently to many aspects of participants' behaviors.

6.5.2 Effects on Performance

The study was able to show that the incremental condition enabled participants to finish the tasks faster. This result is in line with previous works on incremental speech, which have found that incrementality increases performance (Kennington et al., 2014; Skantze & Hjalmarsson, 2010; 2013). Previous works have shown that an incremental system is faster, since it can begin producing responses before the system is done processing what it is going to say. Other work (Kennington et al., 2014) has shown that incremental feedback enables participants to perform better. Incrementality can therefore contribute to efficiency in a system, but also enable people who use such a system to become more efficient in what they do. The current study contributes to evidence of that latter category. In particular, it shows that participants orientating to chunks-sized instructions, such as "to the right", and "a bit more to the right' could more easily find the objects in question than if the robot told them the precise location of the object in one go, by saying for example, "it's on the top shelf".

6.5.3 Effects of Performance on Perception

Analyses of interactions between performative and perceptive metrics revealed an interesting relationship. In particular, they showed that in the non-incremental condition, participants' ratings of the robot are generally unaffected by how long it takes them to find the right objects. In other words, difficulties in finding the objects are not reflected negatively on the robot. This is different for participants in the incremental condition. For those participants difficulties in locating the object are reflected by negative ratings of the degree to which participants thought the robot took them into account, the degree to which they thought the robot responded to their actions, and participants' levels of discomfort. A similar statistical trend that did not reach significance, was also observed between performance and competence for participants in the incremental condition.

What is especially interesting about these relationships is that the ratings of the robot in the non-incremental is almost exactly the same as the ratings for the incremental robot, but only when participants were performing well. This may account for why no significant differences for almost any of the subjective measures could be found. The negative ratings of the robot in the incremental condition, when participants are experiencing difficulties are likely linked to to the robot's communication of the common ground it displays it believes is between itself and the participants. In the incremental condition the robot issues directions to the robot on a moment-by-moment basis, and is informed by participants' non-verbal conduct. That is to say that the robot displays an awareness to participants non-verbal conduct and signals to the participants that these behaviors are part of their common ground. The robot therefore displays knowledge of where the participant is and what he or she is doing, and it displays knowledge of where the object is. By doing this the robot makes an epistemic claim (Heritage, 2012) that it knows more than the participant. A reasonable assumption is that participants hold the robot responsible for performance, expressed in negative ratings when the robot is performing poorly (i.e. not directing participants to an object in a timely manner).

For participants in the non-incremental condition, the robot displays on little awareness of the participants whereabouts or conduct. It merely states where an object can be found. As a result, there is much less common ground between participant and robot, and the common ground is not continually updated as is the case for participants in the incremental condition. The robot does not display knowledge or claim to know more than participants to the same extent as the robot in the *incremental* condition does. Participants have therefore no reason to hold the robot responsible, which is expressed in the relatively stable ratings across different levels of performance (see Figure 6.10 and Figure 6.11)

6.6 Conclusion

This study sought to investigate effects of incremental feedback in an object-finding HRI scenario. The study found that incremental feedback enables participants to perform faster. Adding incremental feedback to a robot's communication design increases the perceived common ground between robot and participants. This also means that when participants perform poorly, they hold the robot responsible, as evidenced by lower ratings in those cases. In order to better balance performance and perception, robots capable of providing incremental feedback should also be programmed with mechanisms, allowing robots to detect and resolve problems, by for example initiating interactional repair, or further grounding relevant contexts.

7. Study 5: Proactivity

7.1 Introduction

Proactive[1] behavior in robots is usually considered a feature that contributes to legibility, and one that leads to tightly coupled HRI. Especially the robot's gaze has been found be able to communicate to people what it is doing and about to do (Mutlu, Shiwa, et al., 2009; Mutlu, Yamaoka, Kanda, Ishiguro, & Hagita, 2009). Proactive gaze can function as a display of contextual awareness, as a display of knowledge (Pitsch, Vollmer, Rohlfing, Fritsch, & Wrede, 2014), and as an display of and awareness of what comes next in the interaction. A robot using proactive gaze thus shows to communication partners that it is attending to a joint plan. Proactive gaze orients to the next step of a joint plan.

There are other gaze models that display an awareness to situational factors. One such model is contingent gaze, whose effects were under investigation in Chapter 3, and which displays an awareness to a communication partners' gaze behavior. Likewise, Chapter 5 investigated a model of gaze in which a robot tracks the position a communication partner's face in a 2D space and makes adjustments to its gaze based on this information. What has not yet been investigated are the effects of a gaze behavior that tracks and adjusts to a robot's own body, rather than that of a communication partner's. Such a gaze behavior displays a different kind of contextual awareness than for example proactive or contingent gaze. Rather than taking external contextual information into account, such as a communication partners' gaze or objects in the work space, the gaze behavior under consideration, in the current chapter, displays an awareness towards its own actions. This self-tracking gaze, or reactive gaze as it will be referred to[2], thus displays a temporal sensitivity to its own actions and behavior, through which it makes its actions more legible to its communication partners.

The aim of the current chapter is to compare the perceptual and performative effects of these two gaze behaviors, proactive and reactive gaze, in a collaborative assembly scenario. The focus for the current chapter is to investigate how displays of awareness of a joint action and a shared plan through contingent robot responses affect interaction and perception.

In the course of the experiment, participants instruct the robot to pick up pieces for

[1]The work presented in this chapter and in the chapter that follows are carried out together with Justus Piater, Özgur Erkennt, and Dadhichi Shukla from Intelligent and Interactive Systems at the University of Innsbruck. Their task was primarily to program the robot, although they also contributed to participant recruitment and running the experiments. The work is partially funded by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 610878, 3rdHAND.

[2]While other forms of gaze, such as contingent gaze, also could be classified as reactive gaze, I reserve the term for a gaze behavior that adjusts to one's own bodily movement.

assembly via gestural actions, such as pointing. Gestural pointing send a signal to communication partners that they should keep attending to the object pointed at (Clark, 2005). Withdrawing from a pointing gesture can thus indicate that the person doing the pointing realizes that the robot has understood what to do. While proactive gaze signals that the robot has understood where to go, reactive gaze signals that the robot is currently engaged in an activity in progress. Gaze cues are generally helpful when they are expected be there (Fischer, Foth, Rohlfing, & Wrede, 2011). However, the problem for HRI is that people do not look enough towards the robot in order to distinguish cues from gaze (Admoni, Dragan, Srinivasa, & Scassellati, 2014). The current chapter therefore investigates how the two displays of understanding, derived from the two gaze signals, affect participants' instructional pointing gestures.

7.2 Related Work

Relevant related work concerns proactive gaze on the one hand and pointing gestures as an instructional resource on the other.

7.2.1 Instructional Gesture: Pointing

Pointing is an indicative act of *directing-to*. For the current chapter, it is understood as the prototypical "index-finger pointing" (Kendon & Versante, 2003). Pointing is a way for communication partners to disambiguate and focus attention on an object of interest (Clark, 2003; Streeck, 2015). 'Object' is here to be interpreted in a broad sense and can refer to physical inanimate objects, but also people (Mondada, 2007) or concepts (de Ruirer, 2000). Clark (2005) argues that a sustained pointing gesture signals to a communication partner that they should keep 'attending' to the object pointed at. His model of *directing-to* pointing gestures consists of three phases; an initiation phase, a maintenance phase, and a termination phase. The initiation phase is the onset of the gesture, the maintenance phase is the gesture proper, and the termination phase is the onset of the withdrawal of the gesture.

Much work in HRI on pointing gestures is concerned with implementing appropriate modeling and evaluating the effects of pointing gestures on robots (e.g. Häring, Eichberg, and André (2012), Hato, Satake, Kanda, Imai, and Hagita (2010)), rather than analyzing how people instruct robots using pointing gestures. Other work also look at how pointing and other gestural actions can be recognized by a robotic system (Raza Abidi, Williams, & Johnston, 2013; Shukla, Erkent, & Piater, 2017). Only few works look at how people employ gestures in collaborative assembly, directed at other people or robots. One such study coded all gestures performed by pairs of people who worked together on an assembly task (Gleeson, MacLean, Haddadi, Croft, & Alcazar, 2013). Interestingly, they find that during the production of gestures (not exclusively pointing gestures) people gaze at each other only rarely.

In summary, previous works suggests that proactive gaze in robots generally has a positive effect on the perceptual and performative metrics. However, it is not yet clear if proactive gaze outperforms other gaze models, and if so, to what extent. Proactive gaze helps, according to the literature, communication partners to signal where an action is going to take place next. Thus, taking Clark's (2005) model of *directing-to* pointing gestures into account, it is logical to assume that proactive gaze in a pick and place task influences the duration of the maintenance phase of a pointing gesture in such as way that people make shorter pointing gestures when their communication partner signals that it has understood what to do by gazing toward the area or object in question.

7.2.2 Proactive Gaze

Gaze is, as noted in previous chapters, a powerful social, but also immensely complex signal. Gaze in interaction between people can be used to "warn, call attention to a misbehavior, bring another person to heel, and otherwise control a situation" (Scheflen & Ashcraft, 1976). Furthermore, human gaze is proactive, rather than reactive (Flanagan & Johansson, 2003; Gredebäck & Falck-Ytter, 2015). That is, people make inferences about what a communication partner is currently doing by gazing to the area where they think their partner will be doing some action. Research on interaction between people finds that unaddressed participants in conversation (e.g. a third party) are able to anticipate the next speaker, which they signal using gaze (Holler & Kendrick, 2015). Research on gaze in joint assembly shows that people look mostly to the tools and the workspace relevant for the task rather than at each other (Fussell, Setlock, & Parker, 2003). In short, gaze is used by people in interaction to signal attention for current and projected actions (among other things).

Gaze as a signaling device has also been investigated in interactions between people and robots to some extent. Sciutti, Bisio, Nori, Metta, Fadiga, Pozzo, and Sandini (2012), for example, suggest to test whether humans gaze proactively to a robot's goal-directed actions. This is followed up by Sciutti, Bisio, Nori, Metta, Fadiga, and Sandini (2012), who report that people gaze in similar ways to robot actions as to human actions. They show that robots' gaze directions are seen as meaningful. Further evidence that people consider robots' gaze to be meaningful is provided by Mutlu, Shiwa, et al. (2009), who find that people were able to determine which conversational role their robot played based on its gaze behavior. Similarly, Boucher et al. (2012) show that people can successfully anticipate a robot's next action based on its gaze and direction and head movement. Pandey et al. (2013) who show that people perceive a robot as more 'aware' and 'supportive' when the robot uses proactive gaze. Mutlu, Yamaoka, et al. (2009) investigate how people playing a game with a robot can infer its attention. They find that when the robot uses gaze cues, people perform significantly better than when not. Similarly, Ivaldi, Anzalone, Rousseau, Sigaud, and Chetouani (2014) find that people respond faster when the robot signals its attention proactively.

However, these results contradict with other findings that show that robots do not evoke reflexive attentional cueing. That is, while humans react to gaze shifts in other humans, they do not respond to gaze shifts produced by robots (Admoni, Bank, Tan, Toneva, & Scassellati, 2011; Meltzoff, Brooks, Shon, & Rao, 2010). Furthermore, Admoni, Hayes, Feil-Seifer, Ullman, and Scassellati (2013) find that people can recognize shorter, more frequent fixations in robots more easily than longer, less frequent gaze duration.

In summary, the current body of anticipatory, or proactive gaze shows positive perceptual and performative effects. However, many of the studies carried out do not compare the effects of proactive gaze with other potential meaningful gaze models (e.g. object- or face tracking, e.g. Ivaldi et al. (2014), Mutlu, Shiwa, et al. (2009), Sciutti, Bisio, Nori, Metta, Fadiga, and Sandini (2012)), or proactive gaze is compared to non-informative gaze conditions, such as static gaze (e.g. Boucher et al. (2012), Mutlu, Yamaoka, et al. (2009)). It is thus an open question whether the positive performative and perceptual effects of proactive gaze reported on above differ significantly from the effects elicited by other forms of informative gaze models.

7.2.3 Reactive Gaze

Very little work has been carried out on reactive gaze in HRI. A few studies have looked at gaze behaviors that are relevant to the current investigation. For example, one study shows that a robot gazing only at its own actions, without taking a communication partner into account, makes it difficult for people to interact with it (Fischer et al., 2015). Likewise, a robot that approaches people by fixating its gaze on participants was found to be perceived as more intelligent and more cooperative than when it fixated its gaze on its path (Fischer et al., 2016).

Other work in HRI deals more generally with reactive versus proactive behaviors, without targeting gaze specifically. For example, Huang, Cakmak, and Mutlu (2015) compare reactive, proactive and adaptive behaviors in a human-robot handover scenario. They find that participants prefer the robot in the reactive condition, but perform better when the robot uses proactive behaviors. Adaptive behaviors, in which the robot signals readiness when participants also are ready to proceed, seem to take the best from both the proactive and reactive conditions.

Although the research on reactive gaze in HRI is minimal, and proactive gaze is usually compared to static gaze, the literature review indicates that proactive gaze should outperform reactive gaze on both perceptual and performative metrics. However, since to date no direct comparisons of these gaze behaviors have been carried it out, their exact relationship remains an open question.

7.2.4 Hypotheses

Based on the current literature, the following hypotheses emerge:

H_1: A robot will, in a collaborative pick and place task, be perceived as more competent when using proactive gaze to signal understanding of gestural instructions, in comparison to other gaze signals.

H_2: Participants instructing a robot in a collaborative pick and place task will produce pointing gestures of a shorter duration when instructing a robot using proactive gaze, in comparison to other gaze signals.

7.3 Method

In order to test the two hypotheses stated above, an experiment using a collaborative robot platform (see Figure 7.1) was designed. With the aid of the robot, naïve participants were tasked with assembling a plastic simple plastic stool for children.

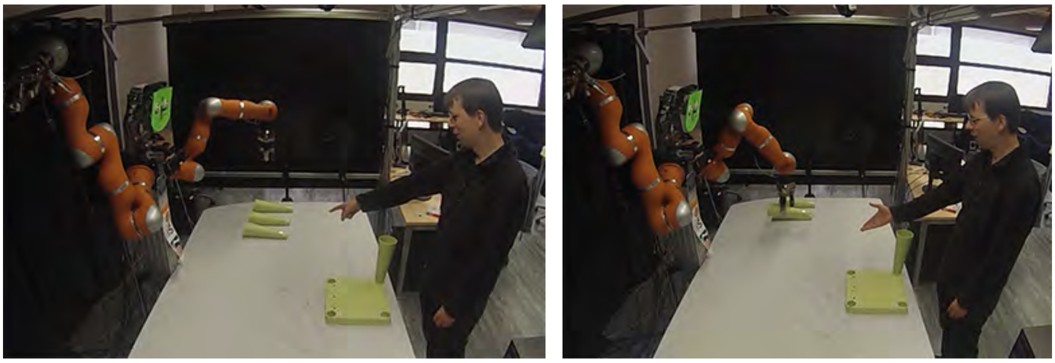
Figure 7.1: The robot

7.3.1 Experimental Conditions

The experiment had two conditions in a between-subjects design. Initially, the robot looked at and tracked participants' faces until it started moving its arm. Previous work on the same platform in similar scenarios has shown that gaze that takes the participant into account positively influences how participants engage with the robot (Fischer et al., 2015). The facial tracking was done using head-pose estimation information acquired from a webcam located just below the robot head. In one condition, the robot gazed proactively. That is, whenever the robot arm moved from one location to another, the robot head indicated where it was going to move to by gazing to this location in the workspace prior to and during robot arm movement, both head pose and eyes fixated on the tartget location. This is referred to as the the *proactive* condition. In the other condition, the robot gazed reactively. That is, whenever the robot's arm moved, the robot 'head' 'followed' the arm via a tracking motion. This is referred to as the *reactive* condition. After each move, the robot face returned to look at and track the face of the participant until it received a new instruction. During the instruction, the experimenter used the two gestures the robot really knows, namely the pointing and handover gestures illustrated in Figure 7.2. However, the experimenter did not draw attention to the gestures, nor did he or she inform the participants about the real conditions of the robot. Instead, he or she explicitly stated that it was up to the them to to figure out how to instruct the robot.

7.3.2 Procedure

Before the experiment, participants were informed that their task would be to assemble an IKEA children's stool with the assistance of the robot. In particular, their task was to instruct the robot to fetch the legs of the stool while the participants had to perform the actual assembly themselves. It was left open to the participants exactly how to instruct the robot. The instruction consisted of two phases: a fetching phase, in which participants had to indicate to the robot which of the four legs they wanted, and a handover phase, in which the participant had to let the robot know where to deliver the leg. The participants then connected the leg to the seat until all four legs were in their respective slots and the stool was assembled.



(a) 'Pick up' gesture (b) 'Hand over' gesture

Figure 7.2: Gestures Recognizable by the Robot

7.3.3 Data

86 participants interacted with an industrial robot to collaborate on the construction of a piece of furniture. 27 were female and 59 were male (age range 18-39), who were distributed equally between the two conditions. Three interactions had to be removed from analysis due to robot breakdowns during the experiments. Interactions lasted about 5 minutes on average. The data analyzed consist of video footage from 2 GoPro cameras and one camcorder placed at different locations in the work space. In addition to the video footage, participants also filled out a demographics questionnaire prior to the experiment, and a questionnaire about participants' perception of the robot after the experiment (for details on the questionnaire, see more below).

7.3.4 Robot

The robot comprises two KuKa arms (Bischoff et al., 2010), each equipped with a Schunk 3-finger gripper, and a KIT head (Asfour, Welke, Azad, Ude, & Dillmann, 2008). However, for this study the robot made only use of its left arm. The robot acted semi-autonomously during the experiments, needing only a confirmation for the planner to execute. Thus, a controller pushes a button whenever the planner is ready to execute. This was to ensure the safety of the participants during the experiments and to avoid damage to the robot.

7.3.5 Protocol

Participants were told what to do, what their role was, and what the role of the robot was, but not exactly how they should instruct the robot. That is, it was left open to participants how they wanted to instruct the robot to pick up and hand over the chair legs to them.

In reality however, the robot only recognized and responded to two different gestures: a pointing gesture for indicating which object it should pick up, and a flat palm gesture for handing over the objects. Previous work on the same platform in similar scenarios has shown that naïve users are more likely to intuitively use gestures before any other form of communication and that the indexical finger-pointing is among the gestures that they are most likely to employ (Jensen, Fischer, Shukla, & Piater, 2015). Based on these results, a classifier was trained to recognize pointing and handover gestures (Shukla et al., 2017). Initially, the robot looks at the participant until it recognizes a pointing gesture in the direction of one of the four stool legs. The classifier had a fairly liberal view of what constitutes 'a pointing gesture' and would also recognize other types of pointing (see Kendon and Versante (2003)) for a review of different types) than the prototypical 'index-finger pointing'. During this 'fetch' phase, the robot used either proactive or reactive gaze, depending on the condition (see subsection on experimental conditions). Once it had picked up an object, it returned its gaze to the participant and awaited further instruction. Once it recognized a flat palm gesture, it moved its arm to a pre-specified location and hand over the object, after which it returned the arm to the 'home' position. Again, the robot gazed according to the experimental condition. Once there, it again looked to the user for the next instruction. This sequence was repeated until all four chair legs had been handed over to the user. In the current study focus is on the first phase (in its four iterations), where participants indicate to the robot with a pointing gesture which object they want the robot to pick up.

7.3.6 Subjective Measures

Apart from demographic information such as, age, sex, level of education and previous experience with robots, participants were asked about their impression of the robot in six questions on a 7-point semantic differential scale. The questionnaire items used here are related to competence in various ways, since H_1 hypothesized that proactive gaze influences participants' perceptions of the robot's competence. The questionnaire items are derived from a slightly modified version of the perceived intelligence scale in the Godspeed questionnaire series (Bartneck et al., 2009). The six items are:

- Incompetent-Competent
- Ignorant-Knowledgeable
- Irresponsible-Responsible
- Unintelligent-Intelligent
- Foolish-Sensible
- Unpredictable-Predictable

In addition, the questionnaire also included three questions about who participants thought was in control and who was responsible for the performance. These were asked as 7-point semantic differentials between robot and participant (1 for participant and 7 for robot). Participant were also asked two additional questions about the collaboration success and on the robot as a collaboration partner, on a 7-point likert scale ranging from not at all (1) to very much (7) ask participants. These questions are, as the previous ones, related to H_1. The underlying assumption is that if the robot is rated more competent as result of the proactive gaze, participants would also to a greater extent see the collaboration as a success.

- How well did you and the robot collaborate?
- To what extent do you consider the collaboration a success?

As a manipulation check, participants were also asked how easy it was for them to guess the robot's next action. This was rated on a 7-point likert scale. Finally, participants were asked when they knew which stool the robot was reaching for. Participants were able to choose one of the following options:

- When it looked at the object
- When it started reaching for the object
- When it opened its hand
- I didn't
- Other (with comment)

Since the robot did not look at the object in the *reactive* condition before the robot arm hovered above the object, the expected distribution of participants in the *reactive* condition should average around "when it opened it's hand", whereas the distribution of participants to average around in the *proactive* condition to center in "When it looked at the object".

7.3.7 Objective Measures

The objective measure of the experiment is the duration of each pointing gesture. In particular, H_2 hypothesized that proactive gaze will lead to shorter pointing gestures. The underlying assumption is that participants will only 'maintain' (Clark, 2005) the pointing until they realize that the robot has understood and is currently processing their request. In order to facilitate such an analysis and to make it quantifiable, participants' pointing gestures and the robot arm movement from the home position to the target location (i.e. above one of the chair legs) during the 'fetch' phases of the experiment were manually coded. Using these two measures, the overlap between participants' pointing gestures were extracted (see Figure 7.3). The overlap measure is converted into a percentage of the of the times it takes from the robot initiates arm movement and until it has reached its target. All annotation and coding is done manually in ELAN (Wittenburg et al., 2006).

Figure 7.3: Coding in Elan

7.3.8 Statistical Analysis

All subjective measures are processed statistically using multiple linear regression, with gender, previous experience with robots and the experimental condition as predictor variables, and questionnaire responses as outcome variables. The objective measure is evaluated by a multiple linear mixed model regression, which takes the experimental condition, previous experience with robot, gender, and the order of instruction[3] as fixed factors, while each participant is a random effect (each participant completed four instructions, one for each leg). The outcome variable is the objective measure (overlap in percent). All quantitative analyses are performed using R 3.3.3. (R Core Team, 2017).

7.4 Results

First, I report on results of the questionnaire analysis, then on the behavioral analysis.

7.4.1 Questionnaire: Manipulation Check

Results from the manipulation check shows that indeed more participants in the *proactive* condition realized early on in the interaction that which chair leg the robot was aiming for compared to participants in the *reactive* condition (see Table 7.1). However, the differences are not statistically significant (χ^2=(2, N=83)=3.91, p=0.14).

	Proactive	**Reactive**
When looking..	39.1%	25.0%
When reaching..	39.1%	61.1%
Open hand..	21.8%	13.9%
Total	100%	100%

Table 7.1: Manipulation Check

[3]Order of the four stool legs

7.4.2 Questionnaire: Ratings

Responses to the questionnaire are evaluated using multiple linear regression as described in the methods section. Results of the six adjectives presented on a semantic differential scale (Figure 7.4) reveal no significant differences between condition. Also, no effects are found for participant gender, or participants' previous experience with robots.

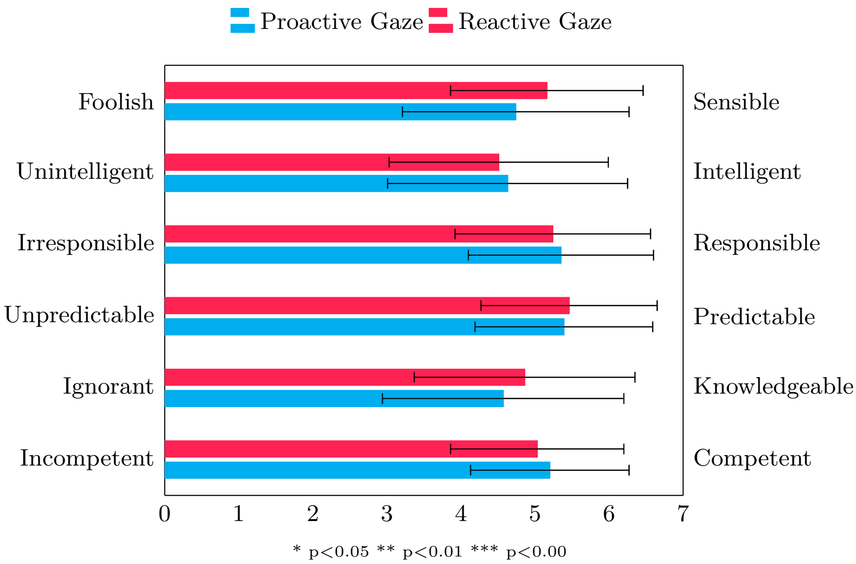
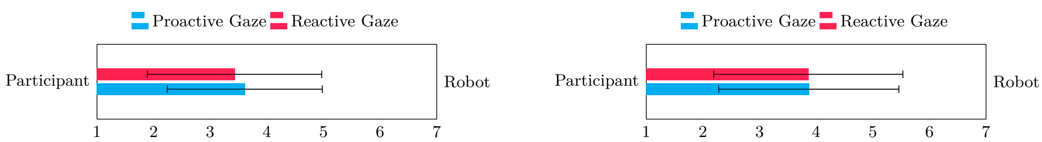


Figure 7.4: Perceived Competence

Likewise, no significant differences are found on the two performance metrics (Figure 7.5).



(a) Responsible for Performance (b) In Control of Performance

Figure 7.5: Performance

Finally, the two Likert Scale items do also show no significant differences between the two gaze conditions.

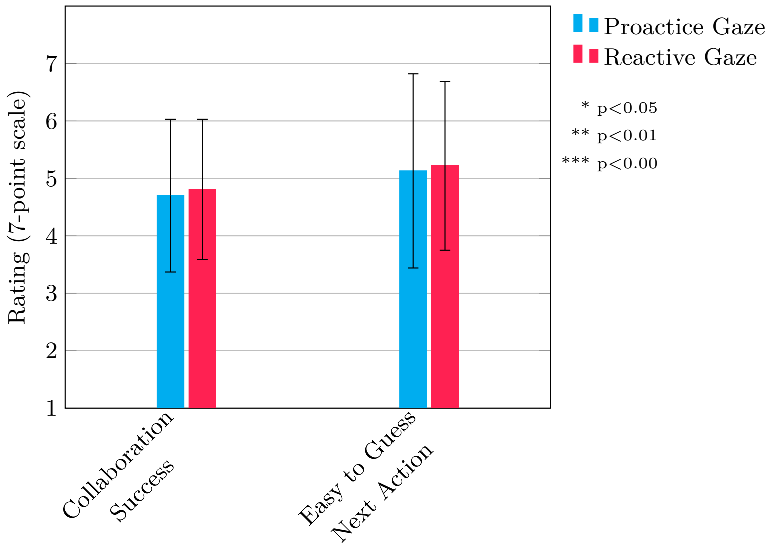


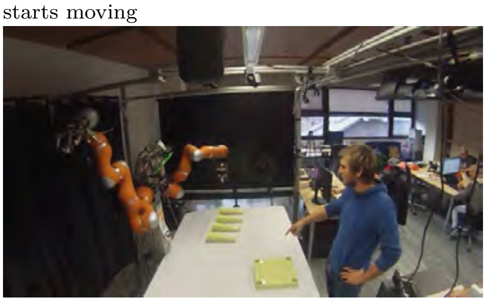
Figure 7.6: Likert Scales

7.4.3 Qualitative Behavioral Analysis

Most participants reduce the duration of their own pointing gestures and the overlap with the robot's movement, over time. That is, with each instruction, participants produce shorter pointing gestures. Since participants are not told how to instruct the robot, this phenomenon is taken as evidence that participants become more proficient in interacting with the robot over time. This is illustrated in the qualitative analysis which serves as an example how participants become more proficient and reduce the duration of their gestures.

1. P points to object

2. P withdraws pointing gesture, as the robots starts moving



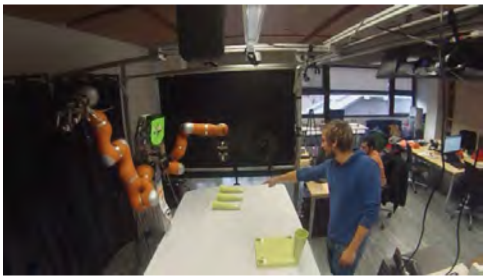
3. P reasserts the pointing gesture

4. P produces other gestures as the robot picks up the leg

Example 7.1: First Instruction

In the first instruction illustrated in Example 7.1, the participant (P) points to the object of interest and maintains the gesture while shifting his gaze between the robot's face and the object until the robot adjusts its gaze and moves its arm. When that happens, he withdraws his gesture, but reengages immediately thereafter. The participants keeps producing gestures, for example grasping and lifting motions, even after the robot has reached the target and is lifting up the object.

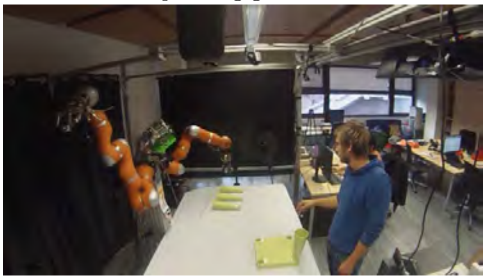
1. P points to object



2. P withdraws pointing gesture, as the robots starts moving

3. P reengages the pointing gesture

4. P withdraws pointing gesture



Example 7.2: Second Instruction

The second instruction illustrated in Example 7.2 is similar to the first one, with the difference that the participant reduces his pointing gesture production. Again, the participant initiates the interaction by producing a pointing gesture in 1 and withdraws it as soon as the robot starts to move. Again, the participant shifts his gaze between the robot's face and the target object between phases 1 and 2. He then reengages the gesture again, but only for a moment, and does not produce any more gestures (pointing or otherwise).

1. P points to object

2. P withdraws pointing gesture, as the robots starts moving

Example 7.3: Third Instruction

In the third instruction illustrated in Example 7.3, the participant initiates with a pointing gesture, which is withdrawn when the robot starts moving and is not reengaged again. Again, several gaze shifts between the robot's face and the target object can be observed.

1. P points to object

2. P withdraws pointing gesture, as the robots starts moving

Example 7.4: Fourth Instruction

As with the prior instructions, the participants initiates the fourth instruction with a pointing gesture, which is withdrawn when the robot starts moving and is not again reengaged. In contrast to the other instructions, the participant keeps his gaze fixated on the robot's gaze for this instruction, until the robot starts moving.

In addition to a potential temporal habituation effect, the current analysis also reveals that participants' gaze behavior is aligned with their gestural instruction. In the examples presented above, the participant displays an orientation to the robot's face as an entry point for the interaction, which is evidenced by his initial gaze toward the robot's face and his return to the robot's gaze whenever he has gazed anywhere else. His gaze shifts to the objects in the workspace can be seen as directives for the robot to follow, in addition to his pointing gesture.

However, not all interactions follow the linear flow described above. The next series of examples shows how both pointing gestures and gaze behavior fluctuate over the four instructions.

1. P points to object

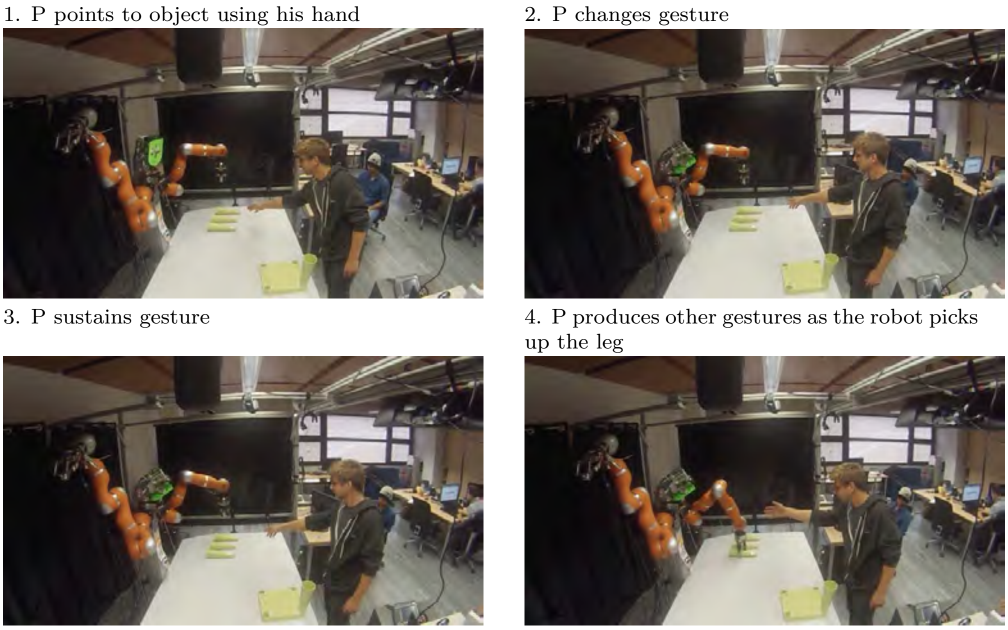
2. P withdraws pointing gesture

3. P points with his hand

4. P changes and sustains pointing gesture

Example 7.5: First Instruction

Initially, the participant points toward an object in the work space, while gazing at the robot face, as also seen in the previous examples. He then withdraws his gesture before he produces another pointing gesture. However, here he points with his entire hand rather than with his index finger. In addition, his gaze is now fixated on the object (leg) in the workspace. When the robot starts moving, he again points towards the target object with his index finger and sustains the gesture until the robot arm has reached the target. His gaze also remains fixed on the target object.

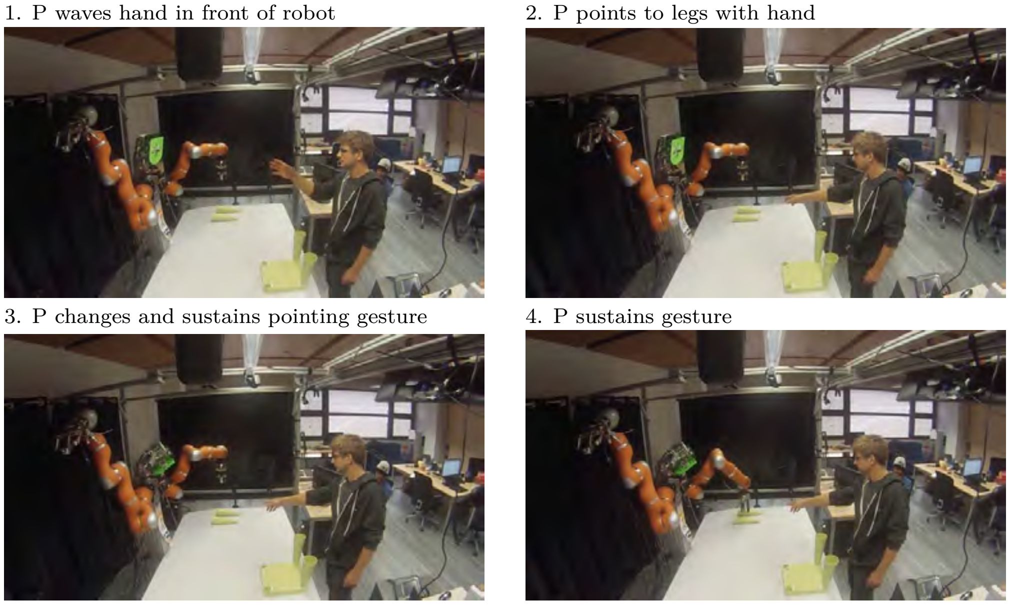


Example 7.6: Second Instruction

For the second instruction, the participant again uses a full hand pointing rather than indexical finger pointing, while looking toward the robot's face. After the robot starts moving he changes the gesture to an index finger point, still keeping his gaze on the robot's face. He sustains this gesture until the robot reaches the target, after which he produces gestural motions as also seen in Example 7.1.

For the third instruction, the participant initially waves at the robot. This can be seen as a pre-sequence (Schegloff, 1980) to the instruction proper. That is, it serves as a call for the robot's attention and indicates that the participant is about to give the robot an instruction. The participant changes the wave into a full hand pointing gesture, which points at the target object in the work space. This is also where the participant has fixated his gaze. The full hand pointing is transformed to an indexical finger pointing, which is sustained until the robot has reached the target object.

For the fourth and last instruction, the participant initially holds up his hand in front of the robot. The motion is not as expressive as the waving he produces in the previous instruction, but it is likely that it performs the same function, which is, to call the robot's attention and to indicate the beginning of an instruction. The motion is transformed into an indexical pointing gesture, while the participant maintains his gaze on the robot's face. He sustains the gesture and his gaze, until the robot starts moving, after which he



Example 7.7: Third Instruction

withdraws his arm.

The analyses of Examples 7.5 to 7.8 show as with the previous examples that over time the participant, through trial and error, does become more proficient at instructing the robot. As illustrated in Examples 7.1 to 7.4 the participant's gaze shifts between the robot and the workspace. This begs the question how participants' gaze behavior affects how they produce gestures and instruct the robot.

7.4.4 Quantitative Analysis: Initial Gaze

The behavioral analysis revealed that participants' gaze behavior during the instruction can differ quite a bit. The manipulations of the robot's gaze is contingent on the assumption that participant notice the robot's gaze behavior. Since previous works have shown that people not necessarily monitor a robot's gaze (Admoni et al., 2011; Fischer et al., 2011), participants' gaze are analyzed in order to ascertain whether that is also the case here.

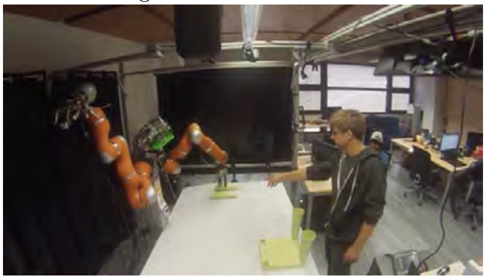
Results of participants' initial gaze when the robot arm starts to move show that 25.4% of the participants in the proactive condition look at the objects in the workspace when the robot starts its gaze action, 63.3% look at the robot's face, while the remaining 38.5% look elsewhere in the workspace, for example towards the parts already assembled. The distribution is almost the same for participants in the reactive condition; 26.3% look at the target objects in the work space right when the robot starts its gaze action, 62.4% look at the robot's face, while 11.3% look elsewhere (χ^2=(1, N=310)=0.03, p=0.98). Interactions in which participants looked to the experimenter (N=7) are dropped from the subsequent analysis.

1. P waves hand in front of robot

2. P points to legs with hand

3. P changes and sustains pointing gesture

4. P sustains gesture



Example 7.8: Fourth Instruction

7.4.5 Quantitative Analysis: All Effects

First, H_2 stated that participants in the *reactive* condition will sustain their pointing gesture for a longer time than participants in the *proactive* condition. This is evaluated in a linear multiple regression analysis with the following variables:

Name	Type	Referent
Experimental Condition	Factor	Reactive Gaze
Experience with Robots	Numeric	-
Gender	Factor	Women
Order	Numeric	-
Gaze	Factor	Workspace

Table 7.2: Regression Variables

However, results show no significant effect of the gaze condition (B=-8.12, SE=7.33, p=0.27). Nevertheless, results do show a strong significant effect of gender (B=23.70, SE=8.25, p=0.005), of previous experience with robots (B=-17.02, SE=4.3, p=0.0001), order of instruction (B=-3.34, SE=1.27, p=0.009), and of gaze to the robot's face (as in encoded in 7.4.4, B=-11.12, SE=4.07, p=0.006).

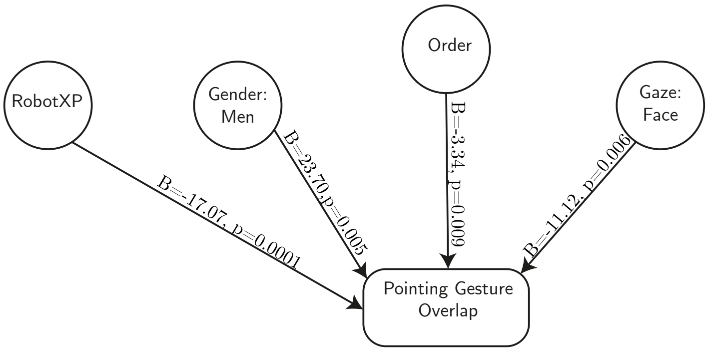
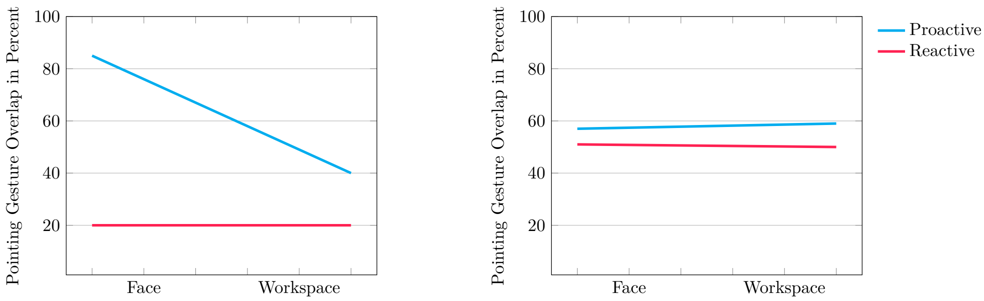


Figure 7.7: Regression Model

Analyses of interactions between variables reveal a marginally significant interaction between participant gender and the experimental condition, and a significant interaction between participant gender and gaze behavior. This relationship is further investigated by separating men and women into two separate data sets, and rerunning the regression analyses on these two new sets of data. Analysis reveals a significant interaction between the experimental condition and where participants initially look, but only for women (see Figure 7.8). Specifically, for women in the proactive condition who looks at the robot's face as it begins its turn condition have a significant lower overlap with the robot's motion than women who look to the workspace (i.e. robot's arm or assembly pieces). However, for women in the reactive condition, their gaze behavior have no effect on their pointing gesture (B=46.22, SE=19, p=0.02). For men no such interaction can be observed. Instead men who look to the robot's face as it begins its turn generally overlap with the robot's motion less than when they look to the workspace.



(a) Women (b) Men

Figure 7.8: Interaction Between Experimental Condition and Gaze

Analyses show that men sustain their gestures significantly longer than women do regardless of condition, and that the duration of the gesture overlap is negatively correlated with previous experience with robots. That is, the more experience with robots participants have, the shorter gestures they produce. Order is correlated negatively with overlap such that the overlap between participants' gestures and the robot's movement decreases with every consecutive instruction. Also, participants who look to the robot's face when it starts moving produce significantly shorter pointing gestures than participants who look to the target objects in the workspace.

7.5 Discussion

The results show no support for neither H_1, which predicted positive ratings for the robot in the *proactive* gaze condition, nor for H_2, which predicted a shorter duration of participants' pointing gestures. There may be several reasons for this. One explanation may be found in the fact that the study rests on the assumption that the robot's gaze behavior in the *proactive* condition is meaningful for participants. Results from other studies (e.g. Mutlu, Shiwa, et al., 2009; Mutlu, Yamaoka, et al., 2009; Sciutti, Bisio, Nori, Metta, Fadiga, Pozzo, and Sandini, 2012; Sciutti, Bisio, Nori, Metta, Fadiga, and Sandini, 2012) show that proactive gaze can be useful. Yet, in this experiment proactive gaze was used to distinguish four objects in very close proximity to each other (see Figure 7.1), which could contribute to the robot's behavior being more ambiguous and less legible. However, if this was in various ways the case, participants should have an easier time interpreting the robot's gaze as more and more objects left the workspace, which would be expressed here as a negative correlation between order and pointing duration overlap. There is some evidence of this. Order was found to correlate negatively with gesture duration. However, the negative correlation applies to participants in both experimental conditions, which indicates that participants become more proficient in interacting with the robot, rather than that the gaze becomes more informative.

7.5.1 Interpersonal Effects on Pointing Gesture Duration

As seen in the results, gender has a substantial impact on the pointing duration to such a degree that it merits a closer inspection. Men's pointing gestures are generally longer than women's. It is worth to note again that more men than women were sampled, so any effects seen here may be exaggerated. A possible explanation could be that men do not look at the robot's face when it's initiating its gaze action as much as the women do. While that is in fact the case (see Table 7.3), this is already accounted for in the regression model (see Figure 7.7) and would also only account for participants in the *proactive* condition (see Table 7.1).

	Face	**Workspace**
Women	82.5%	17.5%
Men	66.2%	33.8%
χ^2=(1, N=275)=4.31, p=0.01)		

Table 7.3: Gender & Participant Gaze Cross-tabulation

Furthermore, analyses of interactions between predictor variables revealed that women who looked to the robot's face at the onset of its turn overlap more than less with the robot's motion, in the proactive condition. This means that participants (in this case female participants) produce longer pointing gestures that overlap more with the robot's motion, when they monitor the robot's gaze. Thus, for these participants the robot's proactive gaze is informative, as participants sustain their pointing gesture in the 'maintenance' phase (Clark, 2005). However, it is not 'informing' in the way it was designed to be. Pointing in the maintenance phase is, according to Clark (2005), a signal that a communication partner should keep attending to the object that is being pointed to. In other words, sustained pointing can be taken as signal that participants believe that the robot 'needs' the sustained instruction to be able to complete its task.

The proactive gaze was designed to signal to participants what the robot was going to do next. As such the robot would displays its awareness towards its own current and future actions, and make visible that it takes these actions as common ground between robot and participant. However, the study clearly shows that this is not how participants take its gaze to mean. In the proactive condition, when the robot's gaze at all is monitored, participants' pointing gestures overlapped even more than when not monitoring the robot's gaze.

The behavioral results together with the results from the questionnaire lead to the conclusion that the gaze cue designed to communicate

8. Study 6: Contingent Repair

8.1 Introduction

This chapter reports on displays of awareness to a specific aspect of participants' communication with a robot in a collaborative assembly scenario[1]. Much previous work on human interaction shows that interaction partners adjust to each other over the course of interactions, and that particularly in repeated tasks they develop shared representations, for instance concerning the lexical material used (Clark & Wilkes-Gibbs, 1986), but also concerning interactional procedures (Mills, 2014), leading to increasingly tight coordination (see also Chapter 7).

Human interaction is characterized by considerable interactional coordination (Clark, 1996; Clark & Krych, 2004; Sacks et al., 1974). In general, interaction partners respond to each other in a time frame of about 300-500 milliseconds, which requires the successful prediction of next actions (Jefferson, 2004; Sacks et al., 1974). Thus, in iterative tasks, with the next action becoming more predictable, interactions between humans become increasingly fluent and tightly coupled. Similar observations have been made for human-robot interaction; for instance, Fischer (2016a) describes how people adjust to robotic communication partners over time.

Another way in which people adjust to each other in interaction is through interaction repair. Repair (see Schegloff (1979)), refers to a set of methods through which participants in interactions can address and resolve interactional trouble. Repair is generally distinguished along three dimensions; initiator of the repair action (self or other), producer of the trouble of source (self or other), and the repair action proper (address or correction). The kind of repair action under investigation for the current chapter is other-initiated self-repair, in which the repair iniator offers a candidate solution to the interactional problem produces by their interaction partner.

This chapter investigates how a robot is able to display its awareness towards certain aspects of participants' communication with it by responding to repair initiated by participants after the robot has made an error. The error made by the robot is not fatal, or even critical, but it may be treated by participant as interaction trouble, which is what Suchman refers to as "communicative breakdowns" (1987). Of special interest is what Suchman refers to as "garden paths". "Garden paths" happen when operational errors occur, to which one party is ignorant of. While the ignorant parties in Suchman's studies are the human users, it is a robot in the current study. That is, the robot produces an error for which it displays

[1] Parts of this chapter are based on Jensen, Fischer, Kirstein, et al. (2017).

no awareness in one condition, while under other it does in another.

More specifically, this chapter investigates a display of contextual awareness, in which a robot is able to change its online behavior based on a human communication partner's gestural action. The robot, which is identical to the one in Chapter 7, hands over pieces of a stool for assembly on the human participant's request. However, during instruction, the robot does not always pick up the right piece. The investigation in this chapter focuses on behavioral and perceptual effects of responses to repair initiations by the participants on the robot's actions.

8.2 Previous Work

Previous work concerns adaptive behaviors and interactional repair in HRI.

8.2.1 Adaptive Behavior in HRI

Much work on adaptive behaviors in HRI focuses on user preference. To little surprise, people generally prefer robots that adapt to the situation and to their behavior than robots that do not (Huang et al., 2015; Sekmen & Challa, 2013).

Other work focuses on behavioral effects, such as efficiency in handover tasks. For example, Huang et al. (2015) find that a robot adapting its timing to humans' behaviors in a handover task increases team performance. Other studies also show that human-robot team performance (usually measured in response time or completion time) increases when robots adapt to human behavior (Hemminghaus & Kopp, 2017; Nikolaidis, Zhu, Hsu, & Srinivasa, 2017).

Adaptability can also be achieved by giving a robot the ability to ask clarification questions. Using such a method, Wu et al. (2015) show how human-robot handovers become more fluent. Broad, Arkin, Ratliff, Howard, and Argall (2017) use a similar approach, however, without testing its effects against a control condition.

Thus, adaptive behaviors are generally considered both perceptually and performance-wise superior to non-adaptive behaviors.

8.2.2 Repair in HRI

Repair is a special kind of adaption to behavior as it arises only when communication partners experience interactional trouble. The interactional trouble under investigation for the current study are errors produced by the robot, to which the human participant initiates repair. Previous work on robots that make errors shows that the robots are generally perceived as less reliable and less trustworthy (Lee, Kiesler, Forlizzi, Srinivasa, & Rybski, 2010; Salem, Lakatos, Amirabdollahian, & Dautenhahn, 2015) and less intelligent (Bajones, Weiss, & Vincze, 2016). However, several works also show that, while the perception of robots is influenced negatively when they make errors, people's behavior towards them is not (Andersen, Köslich, Pedersen, Weigelin, & Jensen, 2017; Mirnig et al., 2017; Robinette, Li, Allen, Howard, & Wagner, 2016; Salem et al., 2015).

There are several ways in which robots can mitigate potential negative effects of errors. Robots can, for example, initiate repair and ask for help (Bajones et al., 2016; Honig & Oron-Gilad, 2018), apologize (Lee et al., 2010; Shiomi, Nakagawa, & Hagita, 2013), or attempt to identify and self-initiate repair (Spexard, Hanheide, Li, & Wrede, 2008). When it comes to repair in HRI, much of the work carried out focuses on repair in verbal interactions. Within this setting, repair initiations by human communication partners has been shown to be an indicator of engagement (Jensen, 2016) and to increase satisfaction (Gonsior, Landsiedel, Glaser, Wollherr, & Buss, 2011). The lack of ability to respond to repair initiations has also been found to lead to "unfavorable long turns" (Opfermann & Pitsch, 2017), which is likely to affect perception. However, other studies also describe how people use non-verbal methods, such as gaze (Muhl & Nagai, 2007), to repair a trouble sources in the interaction with a robot. Opfermann, Pitsch, Yaghoubzadeh, and Kopp (2017) describe how groups of people interacting with a robot signal trouble to each other, using gaze.

Previous works on adaptive behaviors in general, and on repair specifically, in HRI indicate that performative and perceptual metrics should increase in a system that is able to respond to repair initiations over a system that is not.

8.3 Methods

The procedure, participants, protocol and robot are identical to the ones reported on in Chapter 7, so they will not be repeated here.

8.3.1 Experimental Conditions

Figure 8.1: Instructional Gesture

The experiment has two conditions in a between-subjects design. In one condition, the robot is able to change its current actions based on participants' gestural activity. In other words, the robot is able to respond to participants' repair initiations in real time. This feature is implemented only during the pickup phase of the experiment, that is, when participants instruct the robot which peg to pick up (see Figure 8.1 below). This will be referred to as the *repair* condition. However, due to the constraints of the planner there is a considerable latency of 3-4 seconds from when a participant initiates repair to when the robot responds. This creates the effect that the robot responds contingently, but outside of what in human interaction is seen as 'timely'. In the other condition, the robot responds only to the first instruction given by participants and is thus not able to respond to repair initiations. This is referred to as the *no repair* condition.

Repair Example

1. Participant points to leg
2. Robot hovers above a leg
3. Robot moves to the right leg
4. Robot picks up the right leg

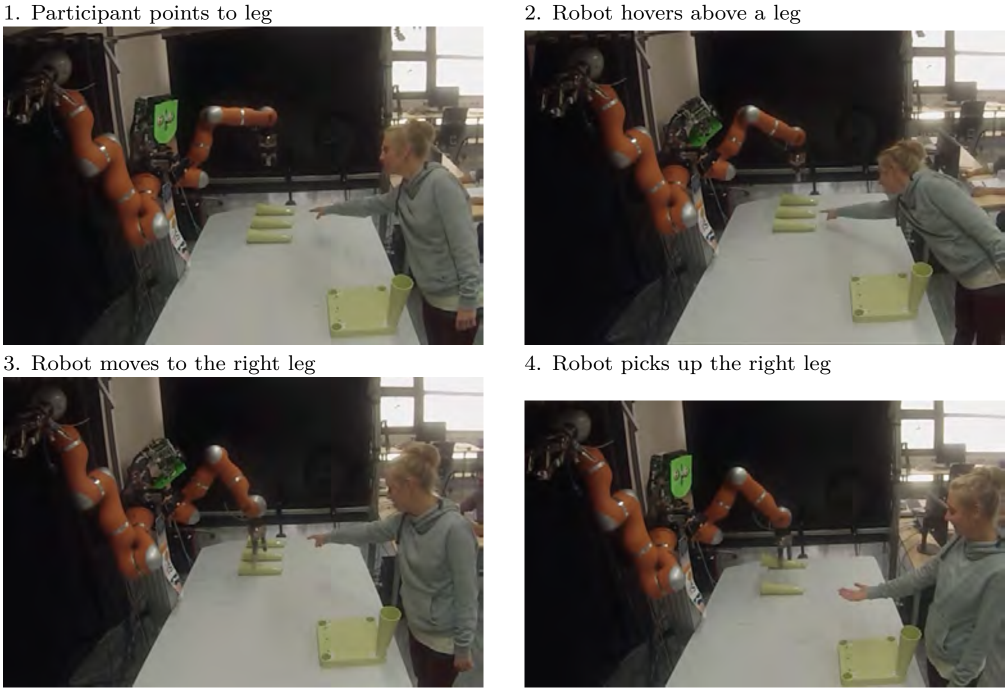


Figure 8.2: Example of Repair

8.3.2 Subjective Measurements

The questionnaire is identical to the one presented in Chapter 7. However, rather than asking when participants realize what the robot is going to do, participants were asked "how easy was it to correct a mistake?" as a manipulation check. The question is presented on a 7-point Likert scale ranging from not at all (1) to very much (7).

8.3.3 Behavioral Analysis

While all subjective measurements are hypotheses-driven (see below), only one aspect of the behavioral analysis is. This aspect is the extent to which the response to repair affects how people instruct the robot in subsequent interactions. A logical alternative to gestural instruction is to speak to the robot. Therefore, whether participants attempt to speak to the robot (to which it did not respond in any condition) was coded for each instruction. In order to ensure that the differences seen between conditions are due the manipulation, rather than interpersonal differences, as for example preference, the data are split in two separate data sets. One data set includes all interactions in which the robot has not yet made an error. Interactions in which the robot never makes an error therefore include all four iterations of the instructions (one for each object to be handed over). A second data set includes all instructions that take place after a robot has made an error. Specific interactions in which the robot produce the error are not considered for this analysis.

8.3.4 Hypotheses

The experiment is evaluated in a between-subject design. Based on the current literature, the following hypotheses emerge:

H_1: The ability to respond to repair initiations influences the robot's perceived intelligence positively.

H_2: The ability to respond to repair initiations influences the extent to which participants feel in control of and responsible for the performance positively.

H_3: The ability to respond to repair initiations influences the robot's perceived adaptability positively.

H_4: The ability to respond to repair initiations gives participants a better understanding of how to interact with the robot, which results in fewer participants who use other methods of instruction than pointing after the robot has made an error in the *repair* condition.

8.3.5 Analysis

A statistical analysis is performed on responses from the questionnaire using multiple linear regression, taking into account effects of sex, previous experience with robots and the experimental condition. In addition to the quantitative analysis, a qualitative analysis is carried out, using ethnomethodological conversation analysis. This analysis helps to clarify what is going on in the interaction and how participants make sense of the activity they are currently engaged in. In order to find out how experimental conditions affect the interaction flow and particpants' sensemaking processes, some of the qualitative findings are then quantified using linear or logistic regression. This type of analysis is what Gehle et al. (2017) refers to as "CA with quantification".

8.3.6 Data

The entire data set consists of 86 interactions (as reported on in Chapter 7). However, for the subjective ratings, only interactions in which the robot made an error are included in the analyses. This was determined qualitatively by examining the trajectory of participants' gestures and comparing them with the robot's actions and its recognition (which was logged). That is, I count an error if a participant (as seen in the video) points at one leg whereas the robot (as seen by the log) identifies another leg. The final data set used comprising only interaction in which some error occurred includes 57 interactions, 22 in the *no repair* condition and 35 in the *repair* condition.

8.4 Results

8.4.1 Subjective Measures

Manipulation Check

Result of the manipulation check show that participants indeed notice the manipulation. Specifically, participants in the *repair* condition find it significantly more easy to correct the robot's action by a mean average of 2.6 points on a 7-point scale (B=2.53, SE=0.46, p=0.000001). No effects for previous experience with robots or for gender are found.

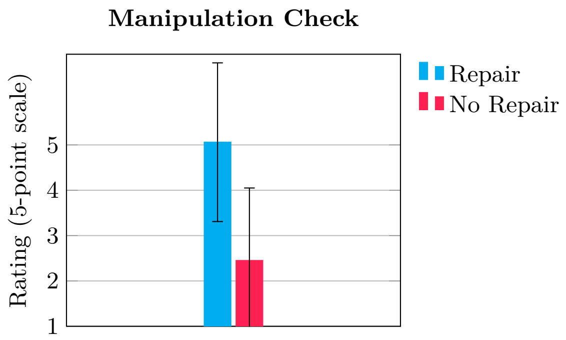


Figure 8.3: How easy was it to correct a mistake?

Questionnaire Results

Analysis of the perceived *intelligence* reveals a significant difference (B=0.69, SE=0.30, p=0.03) on *predictability* (see Figure 8.4). In addition, the analysis also reveals a significant effect on *compliance* (B=.0.74, SE=0.37, p=0.05). That is, participants interacting with the robot in the *repair* condition rate the robot as significantly more *compliant* and *predictable.*

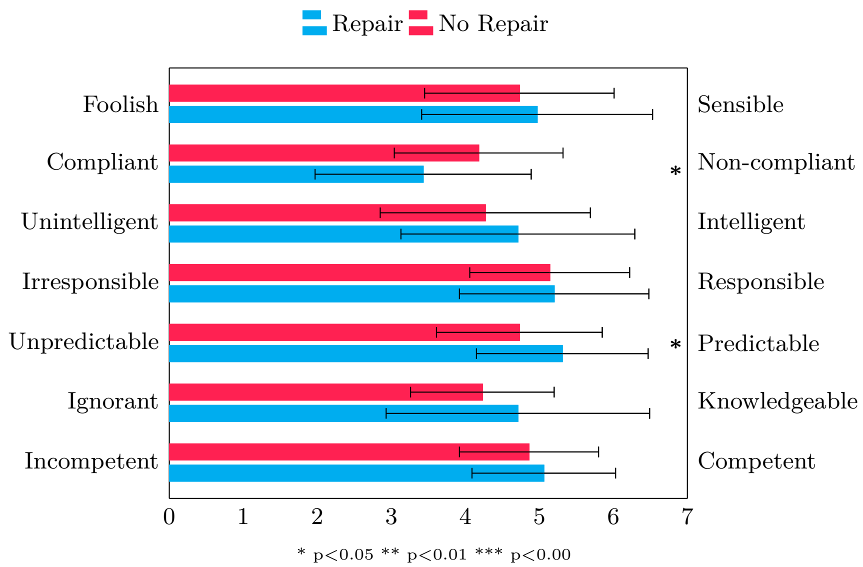
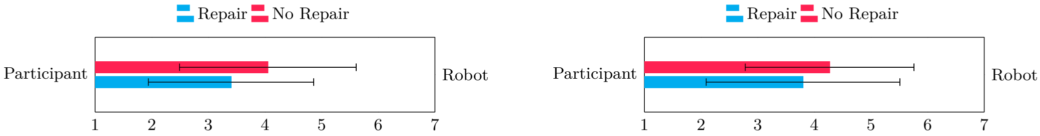


Figure 8.4: Perceived Intelligence

Analysis of who participants thought was in control and responsible for the *performance* of the task shows that, in the *repair* condition participants generally consider themselves more in control (B=-0.53, SE=0.45, p=0.24) and more responsible for the *performance* (B=-0.66, SE=0.41, p=0.11) than participants in the condition in the *no repair* condition. However, the differences are not statistically significant.



(a) Responsible for Performance

(b) In Control of Performance

Figure 8.5: Performance

Analysis of the perceived *adaptability* reveals significant differences between the conditions (see Figure 8.6 below). Specifically, participants found that the robot in the *repair* condition *complied with their instructions* to a greater extent (B=0.84, SE=0.36, p=0.02), and participants thought that the robot was more *eager to collaborate* (B=0.76, SE=0.32, p=0.02) in the *repair* condition than in the *no repair* condition. Analysis also reveals marginal effects of the repair condition on the extent to which participants consider the collaboration successful (B=0.53, SE=0.29, p=0.07). For both of the variables, ratings are more positive in the repair condition.

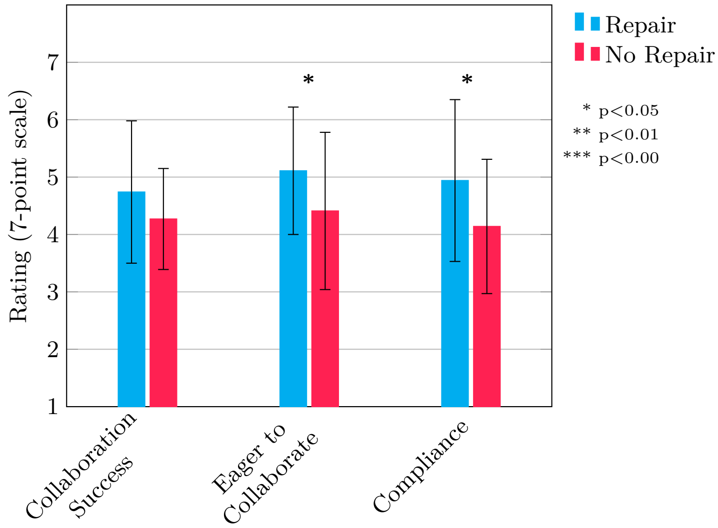


Figure 8.6: Perceived Adaptability

8.4.2 Behavioral Analyses

Analyses show that participants' behavior significantly changes for participants in the *repair* condition after the repair action.

Instruction Phase

The degree to which participants use speech to instruct the robot is evaluated using a logistic regression. Results of participants' instructional behavior show that the likelihood of instructing the robot verbally is 20% in the repair condition and 26.5% in the non-repair condition before the robot has made any error (see Figure 8.7). This difference is not statistical significant (B=-0.38, SE=0.44, p=0.39). No effects for participant gender or previous experience are found.

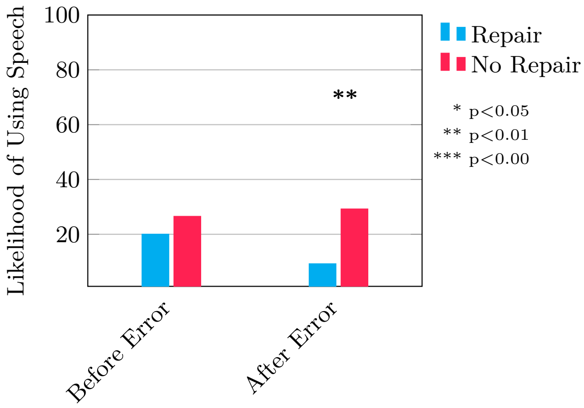


Figure 8.7: Speech to Robot

However, after the robot has made an error, the likelihood of using speech is 26.2% for the *no-repair* condition, while it is only 9.2% for the *repair condition*. This result is statistically significant (B=-1.55, SE=0.56, p=0.006). The likelihood is positively correlated with previous experience with robots, so the more experience, the greater the likelihood that participants will use speech to direct the robot. However, this effect is only marginally significant (B=0.58, SE=0.31, p=0.06) and does not interact with the experimental condition. No effects for participant gender are found.

These results show that the repair action does indeed give participants a better understanding of how to instruct the robot, and in particular how not to do so, evidenced by the vast drop of speech-based instructions from participants in the repair condition after the robot has made an error and participants (in the repair condition) had initiated repair. H_4 is therefore supported.

Handover Phase

Participants are initially unsure as to how to complete the handover phase, which is exemplified in Example 8.1.

Initially (1), the participant waits for the robot and does nothing for 3.1 seconds, until he shifts his gaze to the robot (2), which he sustains for 1.3 seconds. He shifts his gaze again down (3) for another 1.3 seconds, before he again looks at the robot (4), which he sustains for 0.6 seconds. After looking down again (5) for 2.8 seconds, he asks the the robot to hand over the leg while doing a gesture (6). Finally, the robot transports the leg (7), and the participant grabs the leg as the robot releases its grip. The entire sequence takes 20.1 seconds. Each of the gaze shifts displays an orientation, on behalf of the participant, to what he considers to be the robot's turn to perform an action. When this does not happen, he produces a verbal utterance to have the robot hand over the leg. While this utterance has no effect on the robot, the gesture he produces in synchrony with his speech is recognized by the robot, which subsequently initates the handover sequence.

1. Participant waits for the robot to lift the leg

2. Participant looks to the robot face as it stops its motion

3. Participant looks down

4. Participant look up to the robot again

5. Participant looks down again and raises his eyebrows

6. Participant says *"Can you hand it over to me?"* while doing a gesture

7. Participant waits while the robot transports the leg

5. Participant grabs the leg as the robot releases

Example 8.1: First Handover

1. Participant waits for the robot to lift the peg

2. Participant looks to the robot face as it stops its motion

3. Participant says *“hand it over to me”* and produces a gesture

4. Participant grabs the peg as the robot releases

Example 8.2: Fourth Handover

During the course of the experiment, participants adjust to the robot over time and become increasingly savvy about how to interact with the robot best. Thus, the handover sequence for the fourth leg is much smoother as demonstrated, in Example 8.2

In the fourth handover shown above, the participant waits for the robot to lift the peg (1), looks to the robot (2) for 2.6 seconds, produces a verbal utterance together with a gesture (3) and grabs the peg as the robot releases it’s grip (4). While the hesitation in (2) is quite significant and displays an orientation to what the participants considers to be the robot’s turn to perform an action, the interaction is overall more smooth. In comparison, this sequence takes only 14 seconds to complete.

The analysis shows that over the course of the experiment participants learn how to interact with the robot. In order to capture this phenomemon quantitatively, pauses that occur between when the robot has picked up a leg and is ready for the next command and when participants initiate the first handover action are measured. These pauses decrease in duration linearly over the four handover iterations (see Figure 8.8 below).

However, the handover pauses display a high level of interpersonal variability, as shown by the large standard deviations in Figure 8.8. Further analysis shows that for about one third of the participants the interaction does not develop linearly (as depicted in Figure 8.8 and Examples 8.1 and 8.2). While these participants indeed adjust to the robot over time and become increasingly savvy about how to interact with the robot best, they are less fluent in the second execution of the task than they are in the first (see Figure 8.9).

Initially, when the robot stops after it has lifted the first leg, participants initiate the next action after a short delay, which indicates that they cannot predict the robot’s next action (as seen in Example 8.1). However, in round 2, they hesitate even longer, indicating

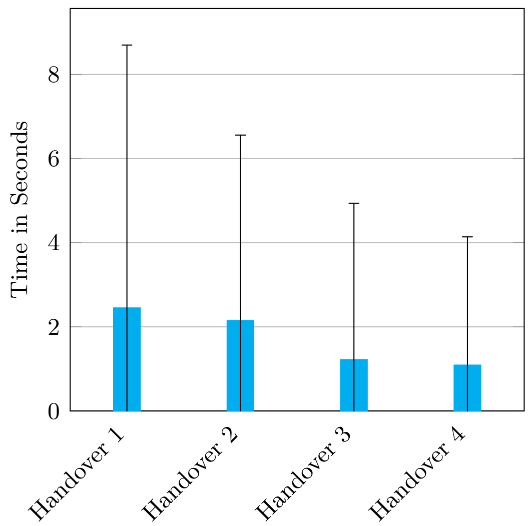


Figure 8.8: Handover Pauses

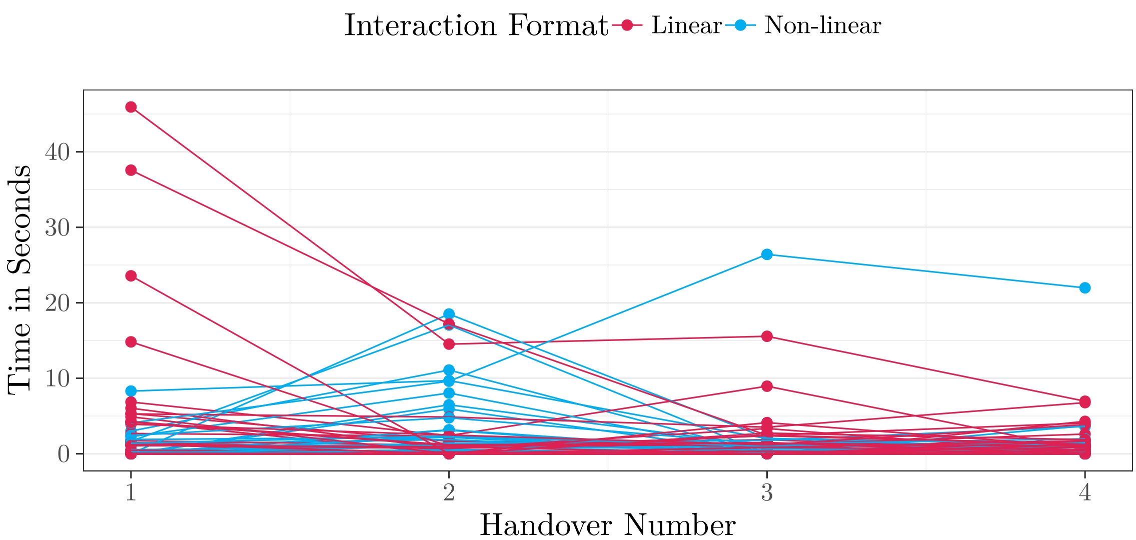


Figure 8.9: Handover Performance

that they expect the robot to carry out its task autonomously. This is demonstrated in Examples 8.3 and 8.4.

The participant first waits for the robot to pick up the leg (1), during which time the participant keeps his gaze fixated on the robot's face. This gaze continues until one second after the robot has picked up the leg. Next, the participant looks down (2) for one second, after which he glances quickly toward the robot's face (3), and then back down again(4), before holding out his own hand in front of the robot (5). He sustains the gesture for 1 second, before he makes a second circular gesture with the same hand (6). Next, he drops his hand down again and waits for 2.1 seconds (7) until the robot starts moving. Finally, he grabs the leg before the robot releases its grip (8). The entire sequence lasts 12.4 seconds. In contrast to Examples 8.1 and 8.2 in which the participant became more fluent in the interaction with each new handover phase, this participant (as well as many others) is less fluent in the second interaction than in the first.

As with the first handover, the participant in the second handover looks towards the robot's face as it finishes its motion (1). He keeps his gaze fixated on the robot's face for 1.4 seconds, whereafter he looks down toward the green leg and (2) and keeps his gaze there for 1 second. He then looks back up, makes a circular gesture (3), looks back down (4), and looks back up the the robot's face (5). These gaze shifts are rapid and last for less than 0.3 seconds. Next, the participant looks down toward the leg again and lets his hand drop down (6). He keeps this posture for 3.1 seconds. He then holds out his arm (7) for 0.4 seconds and againg produces a circular gesture before the robot starts moving and he is able to grasp the leg before the robot releases its hold. The entire sequence is 19.4 seconds long. Thus, rather than becoming more fluent, the participant in Examples 8.3 and 8.4 displays more signs of confusion (for evidenced by, for example, over time the numerous and rapid gaze shifts) Moreover, pauses become longer, not shorter.

The longer stretches of inactivity featured in the beginning of both interactions, as the robot finishes its motion and looks to the participant for the next instruction, are in themselves interesting to investigate further. Initially, when the robot stops after it has lifted the first leg, the participant initiates the next action after a short delay (1 second), which indicates that he cannot predict the robot's next action. However, in handover 2, he hesitates even longer, indicating that he expects the robot to carry out its task autonomously. Thus, the participant assumes that the robot understands that the current task is a repetition of the previous one and that it has successfully learned from the previous interaction what the next step will be, namely to hand over the leg after it has picked it up, without being explicitly signalled to do so again. However, it does become apparant during the second handover that this is not the case and participants are indeed able to recover from the unfulfilled expectation that the robot learns from interaction, as for example people do. Over the course of the experiment almost all participants become more fluent in their interaction with the robot, as indicated by Figure 8.8 and by Example 8.5.

In order to validate the above findings quantitatively, and to find out whether these results interact with the experimental condition, these two interaction formats were coded as either 0 (linear) or 1 (non-linear, i.e. the second handover is longer than the first), and

1. Participants looks to the robots face as it finishes its motion

2. Participant looks down

3. Participant looks back up towards the robots face

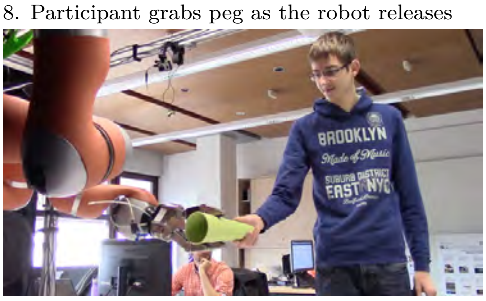
4. Participant looks down again

5. Participants looks toward the robot's face and holds out his hand

6. Participant makes circular gesture with his hand

7. Participant drops his hand again

8. Participant grabs peg as the robot releases



Example 8.3: First Handover

1. Participants looks to the robots face as it finishes its motion

2. Participant looks down toward the peg

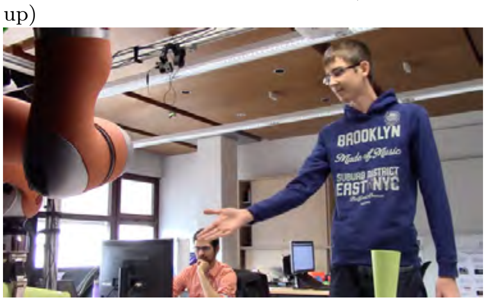
3. Participant looks back up towards the robot's face and makes a circular gesture with his hand

4. Participant looks down again (still doing the circular gesture)

5. Participants looks toward the robot's face

6. Participant looks down again towards the peg and drops his hand down

7. Participant holds out his arm (face palm up)



8. Participant looks up toward the robot's face and makes a circular gesture

Example 8.4: Second Handover

	First Handover	
1.	Robot:	lifts arm with stool leg
2.	*Pause*	(8.30)
3.	Human:	holds out hand
	Second Handover	
4.	Robot:	lifts arm with stool leg
5.	*Pause*	(9.69)
6.	Human:	holds out hand (waiting for robot)
	Third Handover	
7.	Robot:	lifts arm with stool leg
8.	*Pause*	(2.0)
9.	Human:	holds out hand (waiting for robot)
	Fourth Handover	
10.	Robot:	lifts arm with stool leg
11.	*Pause*	(1.27)
12.	Human:	holds out hand (waiting for robot)

Example 8.5: Handover Disfluency

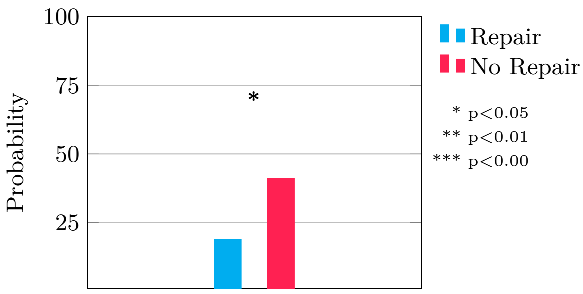


Figure 8.10: Likelihood for Interaction Formats

subsequently used to build a logistic regression model with experimental condition as predictor. The model also includes experience with robots and gender, and the gaze condition as predictors. However, only a significance effect of the repair condition is observed. Results (see Figure 8.10 below) reveal a 40.9% likelihood for the non-linear interaction format for participants in the *non-repair* condition, while the likelihood for the same interaction format for participants in the *repair* condition is just 18.8%. This difference is statistically significant (B=-1.16, SE=0.52, p=0.03). No effects for previous experience with robots or participant gender are found.

8.5 Discussion

The aim of this study was to investigate an aspect of contextual awareness, which concerns a robot's ability to respond to repair initiations by naïve human participants. The study hypothesized that the robot, when responding to repair initiation, would be perceived as more intelligent (H_1), that participants would feel more in control of and responsible for the performance (H_2), and perceive the robot as more adaptable (H_3). Analyses showed support for H_1, H_3, and H_4, while no support could be found for H_2. These findings are very much in line with previous work on adaptability (e.g. Lee et al. (2010), Salem et al. (2015)). The current study shows that the positive effects reported on the ability to be able to respond to repair in spoken interaction e.g. Gonsior et al. (2011), Jensen (2016) between people and robots also applies in situations in which gestural action is the main mode of interaction.

8.5.1 Effects on Perceptual Metrics

Participants found the robot's actions easier to correct, and they found the robot significantly more compliant and predictable in the repair condition, in comparison to the condition without repair. The finding that the robot is rated as more compliant and that its behavior is rated as easier to correct are in themselves not very novel. However, for the vast majority of participants, the robot was able to show off this ability only once, under very specific circumstances, and with a quite substantial delay between repair initiation and response, due to the latency of the robot planner. Even under these tight restrictions, this one action has significant consequences. It indicates that a robot that is able to respond to repair initiations under less restrictive circumstances is likely to affect perceptions of intelligence to an even greater extent.

However, from an HRI perspective it is somewhat surprising that participants rate the robot in the repair condition as more predictable. This result indicates that responses to repair initiations not only contribute to the robot's compliance, but also its legibility. While there are studies that have found a relation between adaptability and legibility (Dehais, Sisbot, Alami, & Causse, 2011; Moon et al., 2014) in human-robot handovers, no studies have to date reported a relation between legibility and responses to repair initiations. However, from a conversation analytical perspective, the result makes a lot of sense. In social exchanges between people, the ability to adjust and ratify the common ground between people is taken for granted. Thus, the response to repair initiation the robot makes in the *repair* condition indicates the robot's awareness toward a certain aspect of a participant's behavior, which works as a very concrete clue as to how to instruct the robot and possibly also how not to do so. This hypothesis is also corroborated by behavioral results, which found that the response to the repair initiation provides participants with a better understanding of how to interact with the robot.

Despite the communicative breakdown created by the robot error, participants generally consider the collaboration as a success, regardless of condition, the difference is although participants in the *repair* condition rate the collaboration as more successful (marginally significant). A possible explanation for this can probably be found in the fact that the error committed by the robot was not fatal. Participants were able to continue the interaction even when the robot did not respond to their repair initiations. Several previous studies posit that one recovery strategy that all robots should have at their disposal is to ignore interactional trouble (Lenz et al., 2012; Opfermann & Pitsch, 2017).

8.5.2 Behavioral Effects

Analyses of participants behavior while interacting with the robot show effects both related to how participants instruct the robot and to how participants coordinate the handover of objects with the robot. The current study shows that the robot's response to repair initiation displays to participants which aspects of their conduct the robot is aware of and responds to.

The study showed that participants in the repair condition indeed changed their instructional behavior after the robot had displayed an awareness toward their repair attempt. Prior to the error, participants in both conditions attempt to control the robot using voice, 20% of the participants in the *repair* condition, compared to 26.2% for participants in the *no-repair* condition. This distribution is comparable to other work on the same robotic platform in a similar scenario (Jensen et al., 2015), which shows that 20% of participants attempt to direct the robot using verbal commands. However, for participants in the repair condition this percentage drops drastically (see Figure 8.7). Participants in the repair condition deal with an objectively more compliant robot, but more importantly the way in which the robot responds to repair indicates to participants how they should interact with it. The robot is responsive to repair and thus adjusts the common ground between them. That is, participants become more aware of what aspects of their communication with the robot it is aware of.

The exploratory analysis revealed two interaction formats when coordinating the handover of objects. In one format the handovers become more fluent over time, which is evidenced by a linear decrease in the time it takes to effect the handovers. However, for about one third of participants a second interaction format was revealed. Here, handovers also become more fluent over time, but in contrast to the first interaction format, the second handover is less fluent than than the first. Further analysis showed that participants in the no-repair condition were significantly more likely to follow the non-linear interaction format than participants in the repair condition. The results show that human-robot collaborations do not simply become more fluent over time, as previous work would suggest; instead, people's expectations that the robot will build on previous interactions results in longer response times and hence less fluent interactions.

People generally bring their experience with interacting with other people to bear in interactions with non-humans, such as robots. Therefore it is logical to assume that after participants have instructed the robot how to do the handover, they expect that it would able to do this autonomously the second time. When that does not happen, interaction trouble surfaces, evidenced by long stretches of non-action. However, at the time of the second handover, participants in the repair condition will already have uncovered some of the limitations of the robot, for example that the robot only responds to gestural instructions, as discussed above. Thus, the common ground between participant and robot for participants in the repair condition are adjusted, so that the participants' partner models of the robot reflect its abilities more accurately. Participants in the *on-repair* condition do not have the same possibilities of adjusting the common ground and are thus more likely to assume that the robot are able to process multiple modalities (as evidenced in Figure 8.7) and learns from instruction (as evidenced in Figure 8.10).

In summary, the studies show that implementing just one opportunity for repair can significantly affect perception and behavior. Specifically, it was shown that the response to repair serves to update participants' partner model of the robot and subsequently changes how they interact with the robot, which methods they use, and how they perceive the robot.

9. Discussion

The overall aim with the six empirical studies in this thesis was to explore how robots' displays of awareness of participants, their behavior, and the context in which the interaction takes place affect interaction and how people perceive robots. Correspondingly, I have studied displays of awareness of the perceptual basis (Chapters 4 and 5), displays of awareness of the actional basis, in particular of proactivity (Chapter 7), contingency (Chapters 3 and 8), incrementality (Chapters 5 and 6), and displays of awareness of the discourse record (Chapter 4). The underlying assumption in all of these studies is that these displays work as signals for what the robot considers common ground. Four of the five indicators were shown to adjust understandings of common ground displayed by participants.

9.1 Indicators for Common Ground

In the following, I discuss how the indicators under investigation affect how participants make their understandings of common ground observable. I relate the discussion to the conceptual model (Figure 1.2) from Chapter 1 and to previous works.

9.1.1 Contingency as an Indicator for Common Ground

Contingency was implemented in three different ways for two different studies. In one study, presented in Chapter 3, a robot was equipped with the ability to produce either contingent gaze or contingent nods. The study showed that contingent gaze contributes significantly to what participants and robot display their common ground to be. Specifically, contingent gaze was shown to enable participants and robot to establish several aspects of common ground. Especially the discourse record and the robot's ability to read the Lego blocks were taken as common ground. Evidence that participants consider the discourse record as common ground is found in the implicit and explicit ways they refer to previous events that had happened in the interaction. Evidence for the robot's suspected ability to read the Lego blocks is found in the way participants structure their instructions for the robot. The establishment of common ground has consequences for what participants perceive the robot is able to do and understand, which in turn affects participants' own verbal and non-verbal behaviors. Other works have previously found that contingent gaze is related to an expectation that robots learn from interaction (Fischer, Lohan, Nehaniv, & Lehmann, 2013; Fischer, Lohan, Saunders, et al., 2013). In other words, people expect a robot using contingent gaze to remember what has happened earlier in the interaction, or what has happened in previous interactions. The results from Chapter 3 add to the body of work that shows that contingent gaze signals an awareness of the discourse record.

In another study, also investigating contingency, Chapter 8, contingency is implemented as a repair mechanism. More specifically, a robot is able to recognize non-verbal repair initiations and respond to them contingently. The study shows that the robot's response to repair initiation, displays to participants which aspects of their conduct the robot is aware of and responds to. The analysis revealed two interaction formats when coordinating the handover of objects. In one format the handovers become more fluent over time, which is evidenced by a linear decrease in the time it takes to effect the handovers, whereas in the other format the second handover takes longer to complete than the first one. Participants in the *no-repair* condition were significantly more likely to follow the non-linear interaction format than participants in the *repair* condition. As I also argue in Chapter 8 that the response to repair initiations makes it clear to participants how they should (and should not) instruct the robot. Thus, this contingent repair grounds the understanding of what the robot is able to understand and what it is able to do.

Thus, for both implementations of common ground, contingency grounds participants' understandings of the robot's ability, which had direct consequences for how people interacted with the robots. Contingent gaze leads participants to make assumptions about the robot's ability to read, understand and recall from earlier on in the interaction. Contingent repair leads participants to reevaluate some of the assumptions they had made of the robot's ability to understand their instruction. It can therefore be said that as contingency signals an awareness of some aspect of an interaction, this signal becomes a cue through which participants understand what the robot understands (or displays to understand).

9.1.2 Incrementality

Incrementality was implemented in two different studies. Incremental feedback had been hypothesized to signal situatedness; it signals to participants that the robot understands and responds to moment-to-moment changes in participants' behaviors. The first study on incrementality (Chapter 5) found that incrementality positively affects intelligence and trustworthiness. Other studies have also found relations between markers of competence, such as intelligence, and incremental feedback (Baumann & Lindner, 2015; Skantze & Hjalmarsson, 2013).

The second study on incrementality (Chapter 6) found only very few differences between *incremental* and *non-incremental* feedback. Surprisingly, the robot in the *incremental* condition was rated as significantly less credible than the robot in the *non-incremental* condition. This stands in contrast to the results from Chapter 5 that found that participants found the incremental robot more trustworthy than the non-incremental robot. This discrepancy is, however, not surprising. Previous works have produced results that seem to contradict each other. For example, incrementality has been reported to positively affect the extent to which participants think a robot is natural (Buschmeier et al., 2012), but other works show the opposite (Chromik et al., 2017). Likewise, one study reports that robots utilizing incremental speech are perceived as more enjoyable (Tsai et al., 2018), while another study reports that a robot utilizing incremental speech is perceived as less likable. For these studies, incremental speech may have been implemented in different ways, which may have had an impact participants' perceptions. However, for the studies in Chapter 5 and Chapter 6, incrementality is implemented in very similar ways, even though the tasks and robots are different. Thus, elements in the makeup of the task could be responsible for the discrepancy. Another explanation could be that because of the difference in task, people perceive the feedback differently. The data at hand does not enable further investigations of this discrepancy. In other words, more (empirical) work needs to be done in order to properly understand the relationship between incremental feedback and perceptive effects, such as credibility and trustworthiness.

The second study on incrementality (Chapter 6) found significant differences in participant behavior between the experimental conditions. Specifically, participants in the *incremental* condition solved their tasks faster than participants in the *non-incremental* condition. This result is in line with previous works on incremental speech, which have found that incrementality increases performance (Kennington et al., 2014; Skantze & Hjalmarsson, 2010; 2013). For participants in the non-incremental condition, the robot displays no awareness of the participants' whereabouts or conduct. It merely states where an object can be found. As a result, there is much less common ground between participant and robot and what common ground there is, is not continually updated as is the case for participants in the incremental condition. Thus, adding incremental feedback to a robot's communication design increases the perceived common ground between robot and participants. Analyses of interactions between performative and perceptive metrics revealed an interesting relationship. In particular, they show that in the non-incremental condition, participants' ratings of the robot are generally unaffected by how long it takes them to find the right objects. In other words, difficulties in finding the objects are not reflected negatively on the robot. This is different for participants in the incremental condition. For those participants difficulties in locating the object are reflected by negative ratings of the degree to which participants thought the robot took them into account, the degree to which they thought the robot responded to their actions, and participants' levels of discomfort.

Interactions between perceptual and behavioral data in Chapter 6 suggest that incremental feedback may increase performance, as previous work also suggests (Chromik et al., 2017; Ghigi et al., 2014; Kennington et al., 2014; Skantze & Hjalmarsson, 2010; 2013), but the increased performance also comes at a price. In other words, participants who interact with a robot that displays an ability to process input and output incrementality hold the robot to a higher standard than participants who interact with a robot without these abilities. Specifically, participants who interacted with the robot endowed with incremental feedback and who for some reason encountered trouble in finding the objects perceived the robot as less competent, more discomforting, felt that to a lesser degree that it responded to their actions and took them into account than participants who either did not encounter any problems or participants who interacted with the robot in the non-incremental condition. This means that these participants hold the robot accountable for their performance. Thus, incrementality signals an awareness to context, which affects the assumptions participants make about the robot. Incrementality grounds participants' understandings of the robot's displays of awareness. This leads to a more smooth interaction when assumptions are met, but can also affect the perception of the robot negatively when problems occur.

9.1.3 Discourse Record

Awareness of the discourse record is implemented in a single study in Chapter 4. The study showed that just a single reference to the discourse record affected the degree to which participants found the robot aware, social and interactive. This reference signals to participants that the discourse record is part of the common ground. Much of previous work encode statistical information (such as game scores) in memory systems for robots (Ahmad et al., 2017; Kipp & Kummert, 2016; Leite et al., 2014). While this work is relevant and important, robots that are designed to engage people in social interaction will also need the ability to be able to encode and interpret talk. Another way in which the results from Chapter 4 differ from previous work is that the robot did not only make a reference to a previous utterance, but recontexualized it for the current utterance. Specifically, the robot used a response to a previous question to be used in a new question where participants are asked to evaluate their experience. While this form of memory access is likely to be more difficult to implement in an autonomous agent than, for example, game statistics, the results from Chapter 4 indicate that the difficulties might be worth the work. The results also resonate with a hypotheses put forth by Christian (2011) who posits that in order for computers (and thus robots) to become more 'human', they need to display that they can make use of and recontextualize past experiences.

9.1.4 The Perceptual Basis

The perceptual basis as an indicator for common ground was implemented in two experiments, presented in Chapter 4 and Chapter 5. Results show that displays of awareness influence perceived traits such as likability and authority, but also influences compliance. Specifically, the displays of awareness of to what happens in the immediate environment support the robot in its other tasks, for example getting participants to drink more water. The perceptual basis is largely unexplored in HRI. Much work on situation awareness in HRI that focuses on a robot's situation awareness, rather than of a controller's (see Chapter 1), investigates participants' actions (e.g. Baxter et al. (2014), Ishii et al. (2013)) rather than the environment participants are in. While this work is relevant and very important, I argue that looking in the immediate space in which interaction takes place may offer equally valuable cues to what an interaction partner might be doing or engaged in. Especially, the study presented in Chapter 5 showed that displays of awareness of the perceptual basis can affect perception and interaction positively. As an indicator for common ground in HRI, the displays of awareness to the perceptual basis signal a joint understanding of physical environment in which interaction takes place.

9.1.5 Proactivity

Proactivity was the only indicator that was not found to contribute to the common ground in human-robot interactions. Proactivity was shown to affect only a subset of participants, and even then these participants did not see the proactive action as an indicator for joint understanding, but rather the opposite. This is not say that proactivity cannot indicate common ground; several works show promising results in this regard (Boucher et al., 2012; Pandey et al., 2013; Sciutti, Bisio, Nori, Metta, Fadiga, Pozzo, & Sandini, 2012; Sciutti, Bisio, Nori, Metta, Fadiga, & Sandini, 2012). However, for Chapter 7 almost no effects regarding proactivity could be found. A possible explanation can be that people simply do not look enough towards the robot because they do not expect to find that level of competence in the robot. Previous studies indicate that this might be the case (Admoni et al., 2014; Fischer et al., 2011). For example, Admoni et al. (2014) had to implement a delay in the handover between their robot and their participants, as participants otherwise would not look at the robot. Whether participants look at the robot or not was explored to some extent in Chapter 7. However, even when participants clearly looked at the robot they did not seem to interpret the gaze signal as a display for understanding or proactive behavior. A possible explanation could be that participants' expectations of the robot are too low. As a direct result thereof the gaze signal produced by the robot does not display an understanding of the joint action, and as such does not contribute to common ground.

9.2 A Model of Common Ground in HRI - Revisited

In the introduction to the thesis, I presented a conceptual model which has guided the investigations I have undertaken. The model is reprinted again in Figure 9.1. The model visualizes that for any given context, participants in interactions display what they take the common ground to be, and that these displays lead to certain effects. What exactly these effects are has been empirically investigated in six chapters, and findings have been discussed and contextualized above.

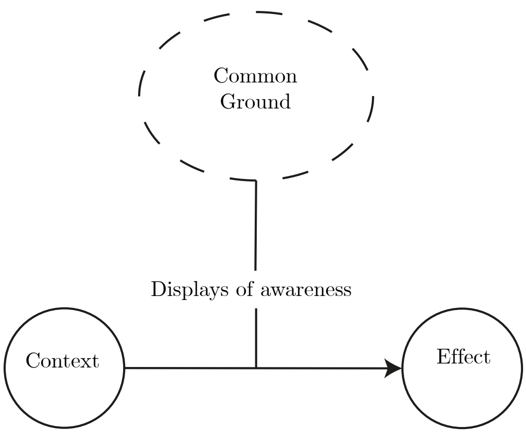


Figure 9.1: Conceptual Model

Throughout this investigation, I have studied how the displays under consideration produced by a robot affect how participants respond to these signals through their behavior or how they perceive of the robot. However, participants' displays of how they understand the robot's signals, are by themselves also a displays of how they understand the common ground. These displays work to adjust perceptions of the common ground. For example, in Chapter 3, contingent gaze was found to affect participants' assumptions of the robot's abilities. Participants interacting with the robot endowed with contingent gaze adjusted their interpretation and understanding of the common ground differently than participants who interacted with the robot without contingent gaze. Specifically, participants' understanding of the robot capabilities expanded in the contingent gaze condition. Similarly, participants' understandings of the robot's capabilities were reduced in the contingent repair condition in Chapter 8.

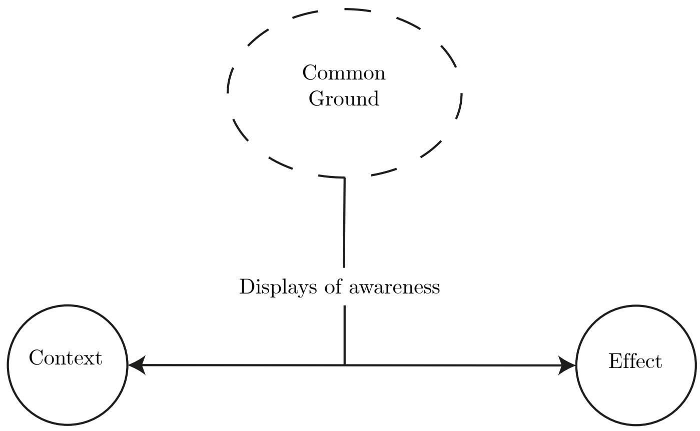


Figure 9.2: Conceptual Model Revised

A conceptual model is slightly modified in Figure 9.2 in order to account for this 'loop', which constantly modifies what participants in interaction understand as common ground. Specifically, the displays, visualized by the arrow, now point in both directions. For interaction between people this is trivial. People adjust their understanding of the common ground on a moment-by-moment basis, or as Clark (1996, p. 92) writes:

> "*[...] we need to keep track of our common ground as it accumulates increment by increment.*"

However, for robots it is a quite different story. The revised model reveals an aspect of interaction that is not often addressed in HRI (if at all). In interaction between people, communication partners can display their understanding of the common ground. This is also what the robots have been programmed to do in this thesis. However, people can also display an understanding of what they understand other people to know, in fact, according to Clark (1996, pp. 110–111) they are quite good at it. Currently, this is not something that robots do. In other words, robots do not display an understanding of participants' understanding, and do not "track" the common ground. This can be potentially problematic for HRI as robots are able to both 'exhibit' and 'claim' understanding (Schegloff, 1982). Exhibiting and claiming understanding that is not there becomes especially problematic when trouble in interaction occurs. For example, in Chapter 6 the robot in the incremental condition displays an awareness to participants' bodily conduct. However, when trouble occurs, the robot is not able to display or to take into account that trouble indeed has occurred. For example, several participants did not know what the word 'napkins' meant, and one participant even went so far as to take out his phone during the experiment in order to look up the word in an online dictionary. Here the robot took the understanding of the word 'napkin' as part of the common ground. While the robot could explain exactly where the napkin was, it had no way of explaining what it takes napkin to mean or even that participants had trouble understanding what it was. While this particular problem could be potentially solved technically in a number of ways, the point is that the robot has no way of updating the common ground, but can only display its understanding of common ground.

In order for robots to engage in social interaction, we as designers, roboticists, and engineers need to find ways in which robots are able to update common ground as interaction unfolds. This can in turn potentially lead to more socially rich and more smooth human-robot interactions.

9.3 Beyond Human-Robot Interaction

Several works indicate that people respond to machines in similar ways as they respond to people (Nass & Moon, 2000; Nass, Steuer, & Tauber, 1994; Reeves & Nass, 1996), for example by assigning agency or applying politeness principles. This phenomenon is generally referred to as Media Equation (Reeves & Nass, 1996). The general idea behind Media Equation is that interaction with media (from a television set to a social robot) is equal to interaction with people. Several works have also found Media Equation to hold for interactions between people and robots (Groom, Chen, Johnson, Kara, & Nass, 2010; Groom et al., 2011b; Groom, Takayama, Ochi, & Nass, 2009). However, in recent years several works have also indicated limitations of Media Equation, especially in relation to Human-Robot Interaction (Bartneck, Rosalia, Menges, & Deckers, 2005; Salem et al., 2013).

The work presented in this thesis has first and foremost implications for human-robot interaction. However, the thesis can also provide a starting point for studying human social interaction in general. For example, in the light of Media Equation, several of the results presented may also be relevant for human social interaction, although it is a claim I cannot make here with the current data at hand.

Studies in conversation analysis show in various ways how social interaction is structured. In particular, they shown the foundational role of contingency in interaction (Schegloff, 1996). Since conversation analytical studies traditionally rely on recordings of interactions that are "naturally occurring" (Hutchby & Wooffitt, 2008, p. 12) it can be difficult to study a specific phenomenon in isolation. In contrast, robots can be programmed to respond contingently to some phenomena, while ignoring others. This opens up for the possibility that HRI studies can also contribute to our understanding of social interaction in general. However, researchers going in this direction will need to find ways to balance the ecological validity between conversation analysis and human-robot interaction.

9.4 Design Implications

In light of the discussion above, several implications for robot design surface.

9.4.1 Design Recommendation I: Incremental Updates to Common Ground

Generally, results show that the more a robot can display its awareness to context, the more favorable it is perceived, and the more seriously it is treated as an interaction partner. However, there is a caveat. The more situationally aware a robot displays itself to be, the more users expect it to be able to perceive, understand and do. This may cause users to overestimate its abilities, which can have problematic consequences for the interaction, as seen in Chapter 6 and discussed above. One way to overcome this problem is by finding ways in which a robot can update its understanding of the common ground shared with a user, for example through repair strategies as investigated in Chapter 8.

9.4.2 Design Recommendation II: Contingency Modifies Perception of Ability

Each of the individual indicators for common ground under investigation also offers perspectives of implications for design. Contingency generally contributes to a feeling of 'situatedness'. This means that robots should rely less on formalized plans and more on responding to cues produced by people. Generally, contingency has been used to expand users' perception of what the robot is able to do and understand (Fischer, Lohan, Nehaniv, & Lehmann, 2013; Fischer, Lohan, Saunders, et al., 2013). However, as Chapter 8 shows, contingency can also be used to indicate a robot's limitations in what it perceives and understands, which eventually leads to smoother HRI.

9.4.3 Design Recommendation III: The Discourse Record

In interactions between people, the discourse record is treated as an interactional resource. Everything that is said and done is always evaluated from the perspective of what has happened before. Robots currently use this resource only to a very small extent. Based on the finding (Chapter 4) that displays of awareness of the discourse record lead to robots being perceived as more social, interactive and aware, one may expect that interaction can be affected as well. In other words, robots displaying this ability may be treated more seriously as interaction partners.

Bibliography

Aarestrup, M., Jensen, L. C., & Fischer, K. (2015). The sound makes the greeting: Interpersonal functions of intonation in human-robot interaction. In *Aaai symposium on turn-taking and coordination in human-machine interaction*.

Abele, A. (1986). Functions of gaze in social interaction: Communication and monitoring. *Journal of Nonverbal Behavior*, *10*(2), 83–101.

Adalgeirsson, S. O., & Breazeal, C. (2010). Mebot: A robotic platform for socially embodied presence. In *Proceedings of the 5th acm/ieee international conference on human-robot interaction* (pp. 15–22). IEEE Press.

Adamides, G., Katsanos, C., Parmet, Y., Christou, G., Xenos, M., Hadzilacos, T., & Edan, Y. (2017). Hri usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer. *Applied ergonomics*, *62*, 237–246.

Adams Jr, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3.

Adams, J. A. (2005). Human-robot interaction design: Understanding user needs and requirements. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 49, *3*, pp. 447–451). SAGE Publications Sage CA: Los Angeles, CA.

Admoni, H., Bank, C., Tan, J., Toneva, M., & Scassellati, B. (2011). Robot gaze does not reflexively cue human attention. In *Proceedings of the cognitive science society* (Vol. 33, *33*).

Admoni, H., Dragan, A., Srinivasa, S. S., & Scassellati, B. (2014). Deliberate delays during robot-to-human handovers improve compliance with gaze communication. Conference Paper. doi:10.1145/2559636.2559682

Admoni, H., Hayes, B., Feil-Seifer, D., Ullman, D., & Scassellati, B. (2013). Are you looking at me?: Perception of robot attention is mediated by gaze type and group size. In *Proceedings of the 8th acm/ieee international conference on human-robot interaction* (pp. 389–396). HRI '13. Tokyo, Japan: IEEE Press. Retrieved from http://dl.acm.org/citation.cfm?id=2447556.2447685

Ahmad, M. I., Mubin, O., & Orlando, J. (2017). Adaptive social robot for sustaining social engagement during long-term children–robot interaction. *International Journal of Human–Computer Interaction*, *33*(12), 943–962.

Ali, M., Alili, S., Warnier, M., & Alami, R. (2009). An architecture supporting proactive robot companion behavior. *In New Frontiers in Human-Robot Interaction at AISB.*

Anastasiou, D., Jokinen, K., & Wilcock, G. (2013). Evaluation of wikitalk–user studies of human-robot interaction. In *International conference on human-computer interaction* (pp. 32–42). Springer.

Andersen, K. E., Köslich, S., Pedersen, B. K. M. K., Weigelin, B. C., & Jensen, L. C. (2017). Do we blindly trust self-driving cars. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 67–68). ACM.

Andrist, S., Tan, X. Z., Gleicher, M., & Mutlu, B. (2014). Conversational gaze aversion for humanlike robots. In *Proceedings of the 2014 acm/ieee int. conference on human-robot interaction* (pp. 25–32). ACM.

Anzalone, S. M., Boucenna, S., Ivaldi, S., & Chetouani, M. (2015). Evaluating the engagement with social robots. *International Journal of Social Robotics*, *7*(4), 465–478.

Aoki, E. (2000). Mexican american ethnicity in biola, ca: An ethnographic account of hard work, family, and religion. *Howard Journal of Communication*, *11*(3), 207–227.

Arend, B., & Sunnen, P. (2017). Coping with turn-taking: Investigating breakdowns in human-robot interaction from a conversation analysis (ca) perspective. In *Proceedings: The 8th international conference on society and information technologies* (pp. 149–154).

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze.* Cambridge U Press.

Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289–304.

Argyle, M., & Graham, J. A. (1976). The central europe experiment: Looking at persons and looking at objects. *Environmental psychology and nonverbal behavior*, *1*(1), 6–16.

Argyle, M., Ingham, R., Alkema, F., & McCallin, M. (1973). The different functions of gaze. *Semiotica*, *7*(1), 19–32.

Asfour, T., Welke, K., Azad, P., Ude, A., & Dillmann, R. (2008). The karlsruhe humanoid head. In *Humanoid robots, 2008. humanoids 2008. 8th ieee-ras international conference on* (pp. 447–453). IEEE.

Asselborn, T., Johal, W., & Dillenbourg, P. (2017). Keep on moving! exploring anthropomorphic effects of motion during idle moments. In *Robot and human interactive communication (ro-man), 2017 26th ieee international symposium on* (pp. 897–902). IEEE.

Bajones, M., Weiss, A., & Vincze, M. (2016). Help, anyone? a user study for modeling robotic behavior to mitigate malfunctions with the help of the user. *arXiv preprint arXiv:1606.02547*.

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.

Bartneck, C., Rosalia, C., Menges, R., & Deckers, I. (2005). Robot abuse—a limitation of the media equation. In *Proceedings of the interact 2005 workshop on agent abuse, rome.*

Bartneck, C., Suzuki, T., Kanda, T., & Nomura, T. (2007). The influence of people's culture and prior experiences with aibo on their attitude towards robots. *Ai & Society*, *21*(1-2), 217–230.

Baumann, T., Kennington, C., Hough, J., & Schlangen, D. (2017). Recognising conversational speech: What an incremental asr should do for a dialogue system and how to get there. In *Dialogues with social robots* (pp. 421–432). Springer.

Baumann, T., & Lindner, F. (2015). Incremental speech production for polite and natural personal-space intrusion. In *International conference on social robotics* (pp. 72–82). Springer.

Baur, T., Damian, I., Lingenfelser, F., Wagner, J., & André, E. (2013). Nova: Automated analysis of nonverbal signals in social interactions. In *International workshop on human behavior understanding* (pp. 160–171). Springer.

Baxter, P., & Belpaeme, T. (2014). Pervasive memory: The future of long-term social hri lies in the past. In *Third international symposium on new frontiers in human-robot interaction at aisb*.

Baxter, P., Kennedy, J., Vollmer, A.-L., de Greeff, J., & Belpaeme, T. (2014). Tracking gaze over time in hri as a proxy for engagement and attribution of social agency. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 126–127). ACM.

Begum, M., Huq, R., Wang, R., & Mihailidis, A. (2015). Collaboration of an assistive robot and older adults with dementia. *Gerontechnology*, *13*(4), 405–419.

Bischoff, R., Kurth, J., Schreiber, G., Koeppe, R., Albu-Schäffer, A., Beyer, A., … Grunwald, G., et al. (2010). The kuka-dlr lightweight robot arm-a new reference platform for robotics research and manufacturing. In *Robotics (isr), 2010 41st international symposium on and 2010 6th german conference on robotics (robotik)* (pp. 1–8). VDE.

Bohus, D., Saw, C. W., & Horvitz, E. (2014). Directions robot: In-the-wild experiences and lessons learned. In *Proceedings of the 2014 international conference on autonomous agents and multi-agent systems* (pp. 637–644). International Foundation for Autonomous Agents and Multiagent Systems.

Boucher, J.-D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., … Ventre-Dominey, J. (2012). I reach faster when i see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in neurorobotics*, *6*, 3.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools.*

Brennan, S. E. (2000). Processes that shape conversation and their implications for computational linguistics. In *Proceedings of the 38th annual meeting on association for computational linguistics* (pp. 1–11). Association for Computational Linguistics.

Broad, A., Arkin, J., Ratliff, N., Howard, T., & Argall, B. (2017). Real-time natural language corrections for assistive robotic manipulators. *The International Journal of Robotics Research*, *36*(5-7), 684–698.

Buschmeier, H., Baumann, T., Dosch, B., Kopp, S., & Schlangen, D. (2012). Combining incremental language generation and incremental speech synthesis for adaptive information presentation. In *Proceedings of the 13th annual meeting of the special interest group on discourse and dialogue* (pp. 295–303). Association for Computational Linguistics.

Cameron, D., Aitken, J., Collins, E., Boorman, L., Fernando, S., McAree, O., ... Law, J. (2015). Framing factors: The importance of context and the individual in understanding trust in human-robot interaction.

Campos, J., & Paiva, A. (2010). May: My memories are yours. In *International conference on intelligent virtual agents* (pp. 406–412). Springer.

Carlmeyer, B., Schlangen, D., & Wrede, B. (2014). Towards closed feedback loops in hri: Integrating inprotk and pamini. In *Proceedings of the 2014 workshop on multimodal, multi-party, real-world human-robot interaction* (pp. 1–6). ACM.

Carlmeyer, B., Schlangen, D., & Wrede, B. (2016a). Exploring self-interruptions as a strategy for regaining the attention of distracted users. In *Proceedings of the 1st workshop on embodied interaction with smart environments* (p. 4). ACM.

Carlmeyer, B., Schlangen, D., & Wrede, B. (2016b). Look at me!: Self-interruptions as attention booster? In *Proceedings of the fourth international conference on human agent interaction* (pp. 221–224). ACM.

Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017). The robotic social attributes scale (rosas): Development and validation. In *Proceedings of the 2017 acm/ieee international conference on human-robot interaction* (pp. 254–262). ACM.

Castellano, G., Aylett, R., Dautenhahn, K., Paiva, A., McOwan, P. W., & Ho, S. (2008). Long-term affect sensitive and socially interactive companions. In *Proceedings of the 4th international workshop on human-computer conversation*.

Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., & McOwan, P. W. (2013). Multimodal affect modeling and recognition for empathic robot companions. *International Journal of Humanoid Robotics*, *10*(01), 1350010.

Castellano, G., Pereira, A., Leite, I., Paiva, A., & McOwan, P. W. (2009). Detecting user engagement with a robot companion using task and social interaction-based features. In *Proceedings of the 2009 international conference on multimodal interfaces* (pp. 119–126). ACM.

Christian, B. (2011). *The most human human: What talking with computers teaches us about what it means to be alive*. Anchor.

Chromik, M., Carlmeyer, B., & Wrede, B. (2017). Ready for the next step?: Investigating the effect of incremental information presentation in an object fetching task. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 95–96). ACM.

Chu, V., Bullard, K., & Thomaz, A. L. (2014). Multimodal real-time contingency detection for hri. In *2014 ieee/rsj international conference on intelligent robots and systems* (pp. 3327–3332). IEEE.

Clark, H. H. (1996). *Using language*. Cambridge University Press Cambridge.

Clark, H. H. (2002). Speaking in time. *Speech Communication*, *36*(1), 5–13.

Clark, H. H. (2003). Pointing and placing. *Pointing: Where language, culture, and cognition meet*, 243–268.

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse studies*, *7*(4-5), 507–525.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition*, *13*(1991), 127–149.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of memory and language*, *50*(1), 62–81.

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, *22*, 245–258.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Cooper, S., Kinsman, L., Buykx, P., McConnell-Henry, T., Endacott, R., & Scholes, J. (2010). Managing the deteriorating patient in a simulated environment: Nursing students' knowledge, skill and situation awareness. *Journal of clinical nursing*, *19*(15-16), 2309–2318.

Côté, N., Canu, A., Bouzid, M., & Mouaddib, A.-I. (2012). Humans-robots sliding collaboration control in complex environments with adjustable autonomy. In *Proceedings of the the 2012 ieee/wic/acm international joint conferences on web intelligence and intelligent agent technology-volume 02* (pp. 146–153). IEEE Computer Society.

Crandall, J. W., & Cummings, M. L. (2007). Developing performance metrics for the supervisory control of multiple robots. In *Human-robot interaction (hri), 2007 2nd acm/ieee international conference on* (pp. 33–40). IEEE.

Cross, E. V. et al. (2009). *Human coordination of robot teams: An empirical study of multimodal interface design* (Doctoral dissertation).

Culpeper, J. (2011). It's not what you said, it's how you said it!": Prosody and impoliteness. *Discursive approaches to politeness*, 57–83.

Cyra, K., & Pitsch, K. (2017). Dealing with 'long turns' produced by users of an assistive system: How missing uptake and recipiency lead to turn increments. In *Robot and human interactive communication (ro-man), 2017 26th ieee international symposium on* (pp. 329–334). IEEE.

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L., & Werry, I. (2005). What is a robot companion-friend, assistant or butler? In *Intelligent robots and systems, 2005.(iros 2005). 2005 ieee/rsj international conference on* (pp. 1192–1197). IEEE.

de Kok, I., Hough, J., Hülsmann, F., Botsch, M., Schlangen, D., & Kopp, S. (2015). A multimodal system for real-time action instruction in motor skill learning. Conference Paper. doi:10.1145/2818346.2820746

de Ruirer, J. P. (2000). 14 the production of gesture and speech. *Language and gesture*, *2*, 284.

Dehais, F., Sisbot, E. A., Alami, R., & Causse, M. (2011). Physiological and subjective evaluation of a human–robot object hand-over task. *Applied ergonomics*, *42*(6), 785–791.

Dickerson, P., Robins, B., & Dautenhahn, K. (2013). Where the action is: A conversation analytic perspective on interaction between a humanoid robot, a co-present adult and a child with an asd. *Interaction Studies*, *14*(2), 297–316.

Dini, A., Murko, C., Yahyanejad, S., Augsdörfer, U., Hofbaur, M., & Paletta, L. (2017). Measurement and prediction of situation awareness in human-robot interaction based on a framework of probabilistic attention. In *Ieee/rsj international conference on intelligent robots and systems (iros)* (pp. 4354–4361). IEEE.

Dole, L. D., Sirkin, D. M., Currano, R. M., Murphy, R. R., & Nass, C. I. (2013). Where to look and who to be: Designing attention and identity for search-and-rescue robots. In *Proceedings of the 8th acm/ieee international conference on human-robot interaction* (pp. 119–120). IEEE Press.

Dominguez, C., Vidulich, M., Vogel, E., & McMillan, G. (1994). *Situation awareness: Papers and annotated bibliography*. Armstrong Laboratory's Situation Awareness Integration.

Drury, J. L., Keyes, B., & Yanco, H. A. (2007). Lassoing hri: Analyzing situation awareness in map-centric and video-centric interfaces. In *Proceedings of the acm/ieee international conference on human-robot interaction* (pp. 279–286). HRI '07. doi:10.1145/1228716.1228754

Drury, J. L., Scholtz, J., Yanco, H. A., et al. (2003). Awareness in human-robot interactions. In *Ieee international conference on systems man and cybernetics* (Vol. 1, pp. 912–918).

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the human factors society annual meeting* (Vol. 32, *2*, pp. 97–101). SAGE Publications Sage CA: Los Angeles, CA.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32–64.

Envarli, I. C., & Adams, J. A. (2005). Task lists for human-multiple robot interaction. In *Robot and human interactive communication, 2005. roman 2005. ieee international workshop on* (pp. 119–124). IEEE.

Fischer, K. (2006). *What computer talk is and isn't: Human-computer conversation as intercultural communication*. AQ-Verlag, Saarbrücken.

Fischer, K. (2016a). *Designing speech for a recipient*. Amsterdam: John Benjamins.

Fischer, K. (2016b). The situatedness of pragmatic acts: Explaining a lamp to a robot. In *Pragmemes and theories of language use* (pp. 901–910). Springer.

Fischer, K., Foth, K., Rohlfing, K., & Wrede, B. (2011). Is talking to a simulated robot like talking to a child? In *Development and learning (icdl), 2011 ieee international conference on* (Vol. 2, pp. 1–6). IEEE.

Fischer, K., Jensen, L. C., Kirstein, F., Stabinger, S., Erkent, Ö., Shukla, D., & Piater, J. H. (2015). The effects of social gaze in human-robot collaborative assembly. In *Icsr* (pp. 204–213).

Fischer, K., Jensen, L. C., Suvei, S.-D., & Bodenhagen, L. (2016). Between legibility and contact: The role of gaze in robot approach. In *Robot and human interactive communication (ro-man), 2016 25th ieee international symposium on* (pp. 646–651). IEEE.

Fischer, K., Lohan, K. S., Nehaniv, C., & Lehmann, H. (2013). Effects of different kinds of robot feedback. In *Social robotics* (pp. 260–269). Springer.

Fischer, K., Lohan, K. S., Saunders, J., Nehaniv, C., Wrede, B., & Rohlfing, K. (2013). The impact of the contingency of robot feedback on hri. In *Collaboration technologies and systems (cts), 2013 international conference on* (pp. 210–217). IEEE.

Fischer, K., Lohan, K., & Foth, K. (2012). Levels of embodiment: Linguistic analyses of factors influencing hri. In *Proceedings of the seventh annual acm/ieee international conference on human-robot interaction* (pp. 463–470). New York, NY: ACM.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769.

Fukuda, H., Kobayashi, Y., Kuno, Y., Yamazaki, A., Ikeda, K., & Yamazaki, K. (2016). Analysis of multi-party human interaction towards a robot mediator. In *Robot and human interactive communication (ro-man), 2016 25th ieee international symposium on* (pp. 17–21). IEEE.

Fussell, S. R., Setlock, L. D., & Parker, E. M. (2003). Where do helpers look?: Gaze targets during collaborative physical tasks. In *Chi'03 extended abstracts on human factors in computing systems* (pp. 768–769). ACM.

Gafaranga, J. (2007). 11. code-switching as a conversational strategy. *Handbook of multilingualism and multilingual communication*, *5*(279), 17.

Garfinkel, H. (1967). *Studies in ethnomethodology.* Englewood Cliffs, NJ, Prentice-Hall.

Gehle, R., Pitsch, K., Dankert, T., & Wrede, S. (2015). Trouble-based group dynamics in real-world hri—reactions on unexpected next moves of a museum guide robot. In *Robot and human interactive communication (ro-man), 2015 24th ieee international symposium on* (pp. 407–412). IEEE.

Gehle, R., Pitsch, K., Dankert, T., & Wrede, S. (2017). How to open an interaction between robot and museum visitor?: Strategies to establish a focused encounter in hri. In *Proceedings of the 2017 acm/ieee international conference on human-robot interaction* (pp. 187–195). ACM.

Gergely, G., & Watson, J. S. (1999). Early socio-emotional development: Contingency perception and the social-biofeedback model. *Early social cognition: Understanding others in the first months of life*, *60*, 101–136.

Ghigi, F., Eskenazi, M., Torres, M. I., & Lee, S. (2014). Incremental dialog processing in a task-oriented dialog. In *Fifteenth annual conference of the international speech communication association.*

Gleeson, B., MacLean, K., Haddadi, A., Croft, E., & Alcazar, J. (2013). Gestures for industry: Intuitive human-robot communication from human observation. In *Proceedings of the 8th acm/ieee international conference on human-robot interaction* (pp. 349–356). IEEE Press.

Goffman, E. (1961). Encounters: Two studies in the sociology of interaction.

Gold, K., & Scassellati, B. (2006). Learning acceptable windows of contingency. *Connection Science*, *18*(2), 217–228.

Gómez, A. V. (2010). *Evolutionary design of human-robot interfaces for teaming humans and mobile robots in exploration missions* (Doctoral dissertation, Universidad Politécnica de Madrid).

Gonsior, B., Landsiedel, C., Glaser, A., Wollherr, D., & Buss, M. (2011). Dialog strategies for handling miscommunication in task-related hri. In *Ro-man, 2011 ieee* (pp. 369–375). IEEE.

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. *Everyday language: Studies in ethnomethodology*, 97–121.

Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological inquiry*, *50*(3-4), 272–302.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of pragmatics*, *32*(10), 1489–1522.

Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., van der Werf, R. J., & Morency, L.-P. (2006). Virtual rapport. In *International workshop on intelligent virtual agents* (pp. 14–27). Springer.

Gredebäck, G., & Falck-Ytter, T. (2015). Eye movements during action observation. *Perspectives on Psychological Science*, *10*(5), 591–598.

Groom, V., Chen, J., Johnson, T., Kara, F. A., & Nass, C. (2010). Critic, compatriot, or chump?: Responses to robot blame attribution. In *Proceedings of the 5th acm/ieee international conference on human-robot interaction* (pp. 211–218). IEEE Press.

Groom, V., Srinivasan, V., Bethel, C. L., Murphy, R., Dole, L., & Nass, C. (2011a). Responses to robot social roles and social role framing. In *International conference on collaboration technologies and systems (cts)* (pp. 194–203). IEEE.

Groom, V., Srinivasan, V., Bethel, C. L., Murphy, R., Dole, L., & Nass, C. (2011b). Responses to robot social roles and social role framing. In *International conference on collaboration technologies and systems (cts)* (pp. 194–203). IEEE.

Groom, V., Takayama, L., Ochi, P., & Nass, C. (2009). I am my robot: The impact of robot-building and robot form on operators. In *Human-robot interaction (hri), 2009 4th acm/ieee international conference on* (pp. 31–36). IEEE.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.

Hanheide, S., Marc and, Lang, C., & Sagerer, G. (2008). Who am i talking with? a face memory for social robots. In *Robotics and automation, 2008. icra 2008. ieee international conference on* (pp. 3660–3665). IEEE.

Häring, M., Eichberg, J., & André, E. (2012). Studies on grounding with gaze and pointing gestures in human-robot-interaction. In *International conference on social robotics* (pp. 378–387). Springer.

Hato, Y., Satake, S., Kanda, T., Imai, M., & Hagita, N. (2010). Pointing to space: Modeling of deictic interaction referring to regions. In *Proceedings of the 5th acm/ieee international conference on human-robot interaction* (pp. 301–308). IEEE Press.

Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources-ca studies of social conduct. *Journal of Pragmatics*, *65*, 1–9.

Hedayati, H., Walker, M., & Szafir, D. (2018). Improving collocated robot teleoperation with augmented reality. In *Proceedings of the 2018 acm/ieee international conference on human-robot interaction* (pp. 78–86). ACM.

Hemminghaus, J., & Kopp, S. (2017). Towards adaptive social behavior generation for assistive robots using reinforcement learning. In *Proceedings of the 2017 acm/ieee* *international conference on human-robot interaction* (pp. 332–340). HRI '17. doi:10.1145/2909824.3020217

Heritage, J. C. (1990). International accountability: A conversation analytic perspective. *Réseaux. Communication-Technologie-Société*, *8*(1), 23–49.

Heritage, J. C. (2012). The epistemic engine: Sequence organization and territories of knowledge. *Research on Language & Social Interaction*, *45*(1), 30–52.

Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: Optimizing recipiency. *Frontiers in Psychology*, *6*(98).

Holliday, A. (1999). Small cultures. *Applied linguistics*, *20*(2), 237–264.

Holmes, J. (1989). Sex differences and apologies: One aspect of communicative competence1. *Applied linguistics*, *10*(2), 194–213.

Holmes, J. (2008). *An introduction to sociolinguistics*. Pearson Longman.

Honig, S. S., & Oron-Gilad, T. (2018). Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in psychology*, *9*, 861.

Hough, J., de Kok, I., Schlangen, D., & Kopp, S. (2015). Timing and grounding in motor skill coaching interaction: Consequences for the information state. In *Proceedings of the 19th semdial workshop on the semantics and pragmatics of dialogue (godial)*.

Huang, C.-M., Cakmak, M., & Mutlu, B. (2015). Adaptive coordination strategies for human-robot handovers. In *Robotics: Science and systems*.

Huang, C.-M., & Mutlu, B. (2016). Anticipatory robot control for efficient human-robot collaboration. In *Human-robot interaction (hri), 2016 11th acm/ieee international conference on* (pp. 83–90). IEEE.

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis*. Polity Press.

Hymes, D. (1964). Introduction: Toward ethnographies of communication 1. *American anthropologist*, *66*(6_PART2), 1–34.

Ishi, C. T., Liu, C., Ishiguro, H., & Hagita, N. (2010). Head motions during dialogue speech and nod timing control in humanoid robots. Conference Paper. IEEE Press.

Ishii, R., Nakano, Y. I., & Nishida, T. (2013). Gaze awareness in conversational agents: Estimating a user's conversational engagement from eye gaze. *ACM Trans. Interact. Intell. Syst.* *3*(2), 1–25. doi:10.1145/2499474.2499480

Ivaldi, S., Anzalone, S. M., Rousseau, W., Sigaud, O., & Chetouani, M. (2014). Robot initiative in a team learning task increases the rhythm of interaction but not the perceived engagement. *Frontiers in neurorobotics*, *8*.

Jefferson, G. (1989). *Preliminary notes on a possible metric which provides for a'standard maximum'silence of approximately one second in conversation*. Multilingual Matters.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Pragmatics and Beyond New Series*, *125*, 13–34.

Jensen, L. C. (2016). Using language games as a way to investigate interactional engagement in human-robot interaction. In J. Seibt, M. Nørskov, & S. Andersen (Eds.), *What social robots can and should do* (Vol. 290, pp. 76–85). doi:10.3233/978-1-61499-708-5-76

Jensen, L. C., Fischer, K., Kirstein, F., Shukla, D., Erkennt, Ö., & Piater, J. (2017). It gets worse before it gets better: Timing of instructions in close human-robot collaboration.

In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction* (pp. 145–146). ACM.

Jensen, L. C., Fischer, K., Shukla, D., & Piater, J. (2015). Negotiating instruction strategies during robot action demonstration. In *Proceedings of the companion of the 2015 acm/ieee international conference on human-robot interaction* (pp. 143–144). doi:10.1145/2701973.2702036

Jensen, L. C., Fischer, K., Suvei, S.-D., & Bodenhagen, L. (2017). Timing of multimodal robot behaviors during human-robot collaboration. In *Ieee international symposium on robot and human interactive communication*, United States: IEEE.

Johnson, S., Rae, I., Mutlu, B., & Takayama, L. (2015). Can you see me now?: How field of view affects collaboration in robotic telepresence. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 2397–2406). ACM.

Jokinen, K., Harada, K., Nishida, M., & Yamamoto, S. (2010). Turn-alignment using eye-gaze and speech in conversational interaction. In *Interspeech* (pp. 2018–2021).

Jung, M. F., Lee, J. J., DePalma, N., Adalgeirsson, S. O., Hinds, P. J., & Breazeal, C. (2013). Engaging robots: Easing complex human-robot teamwork using backchanneling. In *Proceedings of the 2013 conference on computer supported cooperative work* (pp. 1555–1566). ACM.

Kamp, H., & Reyle, U. (1993). *From discourse to logic. introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Studies in Linguistics and Philosophy. Dordrecht/Boston/London: Kluwer.

Kamp, H., & Roßdeutscher, A. (1992). *Remarks on lexical structure, DRS-construction and lexically driven inferences* (Arbeitspapiere des Sonderforschungsbereichs 340 No. 21). Universität Stuttgart. Sprachtheoretische Grundlagen für die Computerlinguistik.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-computer interaction*, *19*(1), 61–84.

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human–robot interaction. *IEEE Transactions on robotics*, *23*(5), 962–971.

Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., & Hagita, N. (2009). An affective guide robot in a shopping mall. In *Proceedings of the 4th acm/ieee international conference on human robot interaction* (pp. 173–180). ACM.

Kasap, Z., & Magnenat-Thalmann, N. (2012). Building long-term relationships with virtual and robotic characters: The role of remembering. *The Visual Computer*, *28*(1), 87–97.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta psychologica*, *26*, 22–63.

Kendon, A., & Versante, L. (2003). Pointing by hand in neapolitan. *Pointing: where language, culture, and cognition meet*, 109–137.

Kendrick, K. H., & Holler, J. (2017). Gaze direction signals response preference in conversation. *Research on Language and Social Interaction*, *50*(1), 12–32. doi:10.1080/08351813.2017.1262120. eprint: https://doi.org/10.1080/08351813.2017.1262120

Kennington, C., Kousidis, S., Baumann, T., Buschmeier, H., Kopp, S., & Schlangen, D. (2014). Better driving and recall when in-car information presentation uses situationally-aware incremental speech output generation. In *Proceedings of the 6th international conference on automotive user interfaces and interactive vehicular applications* (pp. 1–7). ACM.

Kidd, C. D., & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. In *Intelligent robots and systems, 2008. iros 2008. ieee/rsj international conference on* (pp. 3230–3235). IEEE.

Kipp, A., & Kummert, F. (2016). I know how you performed!: Fostering engagement in a gaming situation using memory of past interaction. In *Proceedings of the fourth international conference on human agent interaction* (pp. 281–288). ACM.

Klamer, T., Allouch, S. B., & Heylen, D. (2010). "adventures of harvey"–use, acceptance of and relationship building with a social robot in a domestic environment. In *International conference on human-robot personal relationship* (pp. 74–82). Springer.

Koay, K. L., Sisbot, E. A., Syrdal, D. S., Walters, M. L., Dautenhahn, K., & Alami, R. (2007). Exploratory study of a robot approaching a person in the context of handing over an object. In *Aaai spring symposium: Multidisciplinary collaboration for socially assistive robotics* (pp. 18–24).

Krogsager, A., Segato, N., & Rehm, M. (2014). Backchannel head nods in danish first meeting encounters with a humanoid robot: The role of physical embodiment. In *International conference on human-computer interaction* (pp. 651–662). Springer.

Kružić, S., Music, J., & Stančić, I. (2017). Influence of human-computer interface elements on performance of teleoperated mobile robot. In *Information and communication technology, electronics and microelectronics (mipro), 2017 40th international convention on* (pp. 1015–1020). IEEE.

Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R., & MacDonald, B. A. (2009). Age and gender factors in user acceptance of healthcare robots. In *Robot and human interactive communication, 2009. ro-man 2009. the 18th ieee international symposium on* (pp. 214–219). IEEE.

Kuzuoka, H., Pitsch, K., Suzuki, Y., Kawaguchi, I., Yamazaki, K., Yamazaki, A., … Heath, C. (2008). Effect of restarts and pauses on achieving a state of mutual orientation between a human and a robot. In *Proceedings of the 2008 acm conference on computer supported cooperative work* (pp. 201–204). ACM.

Lallee, S., Hamann, K., Steinwender, J., Warneken, F., Martienz, U., Barron-Gonzales, H., … Ford Dominey, P. (2013). Cooperative human robot interaction systems: Iv. communication of shared plans with naïve humans using gaze and speech. In *Ieee/rsj international conference on intelligent robots and systems (iros)* (pp. 129–136). doi:10.1109/IROS.2013.6696343

Lee, M. K., Kiesler, S., Forlizzi, J., Srinivasa, S., & Rybski, P. (2010). Gracefully mitigating breakdowns in robotic services. In *Human-robot interaction (hri), 2010 5th acm/ieee international conference on* (pp. 203–210). IEEE.

Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, *6*(3), 329–341.

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: A survey. *International Journal of Social Robotics*, *5*(2), 291–308.

Leite, I., Pereira, A., & Lehman, J. F. (2017). Persistent memory in repeated child-robot conversations. In *Proceedings of the 2017 conference on interaction design and children* (pp. 238–247). ACM.

Lemaignan, S., Garcia, F., Jacq, A., & Dillenbourg, P. (2016). From real-time attention assessment to "with-me-ness" in human-robot interaction. Conference Paper. IEEE Press.

Lenz, A., Lallee, S., Skachek, S., Pipe, A. G., Melhuish, C., & Dominey, P. F. (2012). When shared plans go wrong: From atomic-to composite actions and back. In *Intelligent robots and systems (iros), 2012 ieee/rsj international conference on* (pp. 4321–4326). IEEE.

Levinson, S. C. (2016). Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences*, *20*(1), 6–14.

Leyzberg, D., Spaulding, S., & Scassellati, B. (2014). Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 423–430). ACM.

Liddicoat, A. J. (2004). The projectability of turn constructional units and the role of prediction in listening. *Discourse Studies*, *6*(4), 449–469.

Lindwall, O., & Ekström, A. (2012). Instruction-in-interaction: The teaching and learning of a manual skill. *Human Studies*, *35*(1), 27–49.

Linssen, J., Berkhoff, M., Bode, M., Rens, E., Theune, M., & Wiltenburg, D. (2017). You can leave your head on. In *International conference on intelligent virtual agents* (pp. 251–254). Springer.

Liu, C., Ishi, C. T., Ishiguro, H., & Hagita, N. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. Conference Paper. doi:10.1145/2157689.2157797

Lohan, K. S., Pitsch, K., Rohlfing, K. J., Fischer, K., Saunders, J., Lehmann, H., … Wrede, B. (2011). Contingency allows the robot to spot the tutor and to learn from interaction. In *Development and learning (icdl), 2011 ieee international conference on* (Vol. 2, pp. 1–8). IEEE.

Lohan, K. S., Rohlfing, K. J., Pitsch, K., Saunders, J., Lehmann, H., Nehaniv, C. L., … Wrede, B. (2012). Tutor spotter: Proposing a feature set and evaluating it in a robotic system. *International Journal of Social Robotics*, *4*(2), 131–146.

Lubold, N., Walker, E., & Pon-Barry, H. (2016). Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion. In *11th acm/ieee international conference on human-robot interaction (hri)* (pp. 255–262). doi:10.1109/HRI.2016.7451760

MacWhinney, B., & Wagner, J. (2010). Transcribing, searching and data sharing: The clan software and the talkbank data repository. *Gesprachsforschung: Online-Zeitschrift zur verbalen Interaktion*, *11*, 154.

Manuvinakurike, R., Paetzel, M., Qu, C., Schlangen, D., & DeVault, D. (2016). Toward incremental dialogue act segmentation in fast-paced interactive dialogue systems. In

Proceedings of the 17th annual meeting of the special interest group on discourse and dialogue (pp. 252–262).

Matheus, C. J., Kokar, M. M., & Baclawski, K. (2003). A core ontology for situation awareness. In *Proceedings of the sixth international conference on information fusion* (Vol. 1, pp. 545–552).

Mavridis, N., Petychakis, M., Tsamakos, A., Toulis, P., Emami, S., Kazmi, W., … Tanoto, A. (2010). Facebots: Steps towards enhanced long-term human-robot interaction by utilizing and publishing online social information. *Paladyn*, *1*(3), 169–178.

McColl, D., Zhang, Z., & Nejat, G. (2011). Human body pose interpretation and classification for social human-robot interaction. *International Journal of Social Robotics*, *3*(3), 313.

Mehlmann, G., Häring, M., Janowski, K., Baur, T., Gebhard, P., & André, E. (2014). Exploring a model of gaze for grounding in multimodal hri. In *Proceedings of the 16th international conference on multimodal interaction* (pp. 247–254). ACM.

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. (2010). "social" robots are psychological agents for infants: A test of gaze following. *Neural networks*, *23*(8-9), 966–972.

Metta, G., Fitzpatrick, P., & Natale, L. (2006). Yarp: Yet another robot platform. *International Journal of Advanced Robotic Systems*, *3*(1), 8.

Mihalyi, R.-G., Pathak, K., Vaskevicius, N., Fromm, T., & Birk, A. (2015). Robust 3d object modeling with a low-cost rgbd-sensor and ar-markers for applications with untrained end-users. *Robotics and Autonomous Systems*, *66*, 1–17.

Mills, G. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, *32*, 158–173.

Mirnig, N., Stollnberger, G., Miksch, M., Stadler, S., Giuliani, M., & Tscheligi, M. (2017). To err is robot: How humans assess and act toward an erroneous social robot. *Frontiers in Robotics and AI*, *4*, 21.

Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse studies*, *9*(2), 194–225.

Mondada, L. (2009a). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, *41*(10), 1977–1997.

Mondada, L. (2009b). The embodied and negotiated production of assessments in instructed actions. *Research on Language and Social Interaction*, *42*(4), 329–361.

Mondémé, C. (2011). Animal as subject matter for social sciences: When linguistics addresses the issue of a dog's "speakership." *Non-humans in social science: Animals, spaces, things*, 87–105.

Moon, A., Troniak, D. M., Gleeson, B., Pan, M. K., Zheng, M., Blumer, B. A., … Croft, E. A. (2014). Meet me where i'm gazing: How shared attention gaze affects human-robot handover timing. Conference Paper. doi:10.1145/2559636.2559656

Moustris, G. P., Mantelos, A. I., & Tzafestas, C. S. (2013). Shared control for motion compensation in robotic beating heart surgery. In *Proc. of the ieee int. conf. on robotics and automation.*

Muhl, C., & Nagai, Y. (2007). Does disturbance discourage people from communicating with a robot? In *Robot and human interactive communication, 2007. ro-man 2007. the 16th ieee international symposium on* (pp. 1137–1142). IEEE.

Muhl, C., Nagai, Y., & Sagerer, G. (2007). On constructing a communicative space in hri. In *Annual conference on artificial intelligence* (pp. 264–278). Springer.

Mutlu, B., Kanda, T., Forlizzi, J., Hodgins, J., & Ishiguro, H. (2012). Conversational gaze mechanisms for humanlike robots. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *1*(2), 12.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proceedings of the 4th acm/ieee international conference on human robot interaction* (pp. 61–68). ACM.

Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th acm/ieee international conference on human robot interaction* (pp. 69–76). HRI '09. doi:10.1145/1514095.1514110

Mykoniatis, K., Angelopoulou, A., Schaefer, K. E., & Hancock, P. A. (2013). Cerberus: The development of an intelligent autonomous face recognizing robot. In *Systems conference (syscon), 2013 ieee international* (pp. 376–380). IEEE.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81–103.

Nass, C., Steuer, & Tauber. (1994). Computers are social actors. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 72–78). ACM.

Nehaniv, C. L., Dautenhahn, K., Kubacki, J., Haegele, M., Parlitz, C., & Alami, R. (2005). A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction. In *Robot and human interactive communication, 2005. roman 2005. ieee international workshop on* (pp. 371–377). IEEE.

Nikolaidis, S., Zhu, Y. X., Hsu, D., & Srinivasa, S. (2017). Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 acm/ieee international conference on human-robot interaction* (pp. 294–302). HRI '17. doi:10.1145/2909824.3020252

Nonami, K., Shimoi, N., Huang, Q. J., Komizo, D., & Uchida, H. (2000). Development of teleoperated six-legged walking robot for mine detection and mapping of mine field. In *Intelligent robots and systems, 2000.(iros 2000). proceedings. 2000 ieee/rsj international conference on* (Vol. 1, pp. 775–779). IEEE.

Ohshima, N., Fujimori, R., Tokunaga, H., Kaneko, H., & Mukawa, N. (2017). Neut: Design and evaluation of speaker designation behaviors for communication support robot to encourage conversations. In *2017 26th ieee international symposium on robot and human interactive communication (ro-man)* (pp. 1387–1393).

Okuno, Y., Kanda, T., Imai, M., Ishiguro, H., & Hagita, N. (2009). Providing route directions: Design of robot's utterance, gesture, and timing. In *Human-robot interaction (hri), 2009 4th acm/ieee international conference on* (pp. 53–60). IEEE.

Opfermann, C., & Pitsch, K. (2017). Reprompts as error handling strategy in human-agent-dialog? user responses to a system's display of non-understanding. In *2017 26th ieee international symposium on robot and human interactive communication (ro-man)* (pp. 310–316). doi:10.1109/ROMAN.2017.8172319

Opfermann, C., Pitsch, K., Yaghoubzadeh, R., & Kopp, S. (2017). The communicative activity of "making suggestions" as an interactional process: Towards a dialog model for hai. In *Proceedings of the 5th international conference on human agent interaction* (pp. 161–170). HAI '17. doi:10.1145/3125739.3125752

Oto, K., Feng, J., & Imai, M. (2017). Investigating how people deal with silence in a human-robot conversation. In *2017 26th ieee international symposium on robot and human interactive communication (ro-man)* (pp. 195–200).

Pajaziti, A. (2014). Slam–map building and navigation via ros. *International Journal of Intelligent Systems and Applications in Engineering*, *2*(4), 71–75.

Pandey, A. K., Ali, M., & Alami, R. (2013). Towards a task-aware proactive sociable robot based on multi-state perspective-taking. *International Journal of Social Robotics*, *5*(2), 215–236.

Petit, M., Fischer, T., & Demiris, Y. (2016). Lifelong augmentation of multimodal streaming autobiographical memories. *IEEE Transactions on Cognitive and Developmental Systems*, *8*(3), 201–213.

Pillet-Shore, D. (2012). Greeting: Displaying stance through prosodic recipient design. *Research on Language & Social Interaction*, *45*(4), 375–398.

Pitsch, K., & Gehle, R. (2013). Addressing multiple participants: A museum robot's gaze shapes visitor participation. In *International conference on social robotics*.

Pitsch, K., & Koch, B. (2010). How infants perceive the toy robot pleo. an exploratory case study on infant-robot-interaction. In *Second international symposium on new frontiers in human-robot-interaction (aisb)*.

Pitsch, K., Kuzuoka, H., Suzuki, Y., Süssenbach, L., Luff, P., & Heath, C. (2009). "the first five seconds": Contingent stepwise entry into an interaction as a means to secure sustained engagement in hri. In *Robot and human interactive communication, 2009. ro-man 2009. the 18th ieee international symposium on* (pp. 985–991). IEEE.

Pitsch, K., Vollmer, A.-L., & Mühlig, M. (2013). Robot feedback shapes the tutor's presentation: How a robot's online gaze strategies lead to micro-adaptation of the human's conduct. *Interaction Studies*, *14*(2), 268–296.

Pitsch, K., Vollmer, A.-L., Rohlfing, K. J., Fritsch, J., & Wrede, B. (2014). Tutoring in adult-child interaction: On the loop of the tutor's action modification and the recipient's gaze. *Interaction Studies*, *15*(1), 55–98.

Pitsch, K., & Wrede, S. (2014). When a robot orients visitors to an exhibit. referential practices and interactional dynamics in real world hri. In *Robot and human interactive communication, 2014 ro-man: The 23rd ieee international symposium on* (pp. 36–42). IEEE.

Plurkowski, L., Chu, M., & Vinkhuyzen, E. (2011). The implications of interactional repair for human-robot interaction design. In *Proceedings of the 2011 ieee/wic/acm* *international conferences on web intelligence and intelligent agent technology-volume 03* (pp. 61–65). IEEE Computer Society.

Pouget, M., Hueber, T., Bailly, G., & Baumann, T. (2015). Hmm training strategy for incremental speech synthesis. In *Sixteenth annual conference of the international speech communication association*.

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org/

Raza Abidi, S. S., Williams, M., & Johnston, B. (2013). Human pointing as a robot directive. In *Proceedings of the 8th acm/ieee international conference on human-robot interaction* (pp. 67–68). IEEE Press.

Read, R., & Belpaeme, T. (2014). Situational context directs how people affectively interpret robotic non-linguistic utterances. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 41–48). ACM.

Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press.

Richter, V., Carlmeyer, B., Lier, F., Borgsen, S. M. z., Schlangen, D., Kummert, F., … Wrede, B. (2016). Are you talking to me?: Improving the robustness of dialogue systems in a multi party hri scenario by incorporating gaze direction and lip movement of attendees. Conference Paper. doi:10.1145/2974804.2974823

Riek, L. D. (2012). Wizard of oz studies in hri: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1*(1).

Riek, L. D., Paul, P. C., & Robinson, P. (2010). When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, *3*(1-2), 99–108.

Robinette, P., Li, W., Allen, R., Howard, A. M., & Wagner, A. R. (2016). Overtrust of robots in emergency evacuation scenarios. In *The eleventh acm/ieee international conference on human robot interaction* (pp. 101–108). IEEE Press.

Robins, B., Dautenhahn, K., & Dickerson, P. (2009). From isolation to communication: A case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. In *Advances in computer-human interactions, 2009. achi'09. second international conferences on* (pp. 205–211). IEEE.

Rossi, S., Ferland, F., & Tapus, A. (2017). User profiling and behavioral adaptation for hri: A survey. *Pattern Recognition Letters*, *99*, 3–12.

Rossi, S., Staffa, M., Giordano, M., De Gregorio, M., Rossi, A., Tamburro, A., & Vellucci, C. (2015). Robot head movements and human effort in the evaluation of tracking performance. In *Robot and human interactive communication (ro-man), 2015 24th ieee international symposium on* (pp. 791–796). IEEE.

Sacks, H. (1984). Notes on methodology. *Structures of social action: Studies in conversation analysis*, 21–27.

Sacks, H. (1989). Lecture six: The mir membership categorization device. *Human Studies*, 271–281.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 696–735.

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction* (pp. 141–148). ACM.

Salem, M., Ziadee, M., & Sakr, M. (2013). Effects of politeness and interaction context on perception and experience of hri. In *International conference on social robotics* (pp. 531–541). Springer.

Samy Alim, H. (2007). Critical hip-hop language pedagogies: Combat, consciousness, and the cultural politics of communication. *Journal of language, identity, and education*, *6*(2), 161–176.

Sarangi, S. (1994). Intercultural or not? beyond celebration of cultural differences in miscommunication analysis. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *4*(3), 409–427.

Scheflen, A. E., & Ashcraft, N. (1976). Human territories: How we behave in space-time.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation in discourse and syntax. *Syntax and Semantics Ann Arbor, Mich. 12*, 261–286.

Schegloff, E. A. (1967). *The first five seconds: The order of conversational openings* (Unpublished Ph.D dissertation, Department of Sociology, University of California, Berkeley).

Schegloff, E. A. (1968). Sequencing in conversational openings. *American anthropologist*, *70*(6), 1075–1095.

Schegloff, E. A. (1972). Notes on a conversational practice: Formulating place.

Schegloff, E. A. (1980). Preliminaries to preliminaries:"can i ask you a question?" *Sociological inquiry*, *50*(3-4), 104–152.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh'and other things that come between sentences. *Analyzing discourse: Text and talk*, *71*, 93.

Schegloff, E. A. (1987a). Analyzing single episodes of interaction: An exercise in conversation analysis. *Social psychology quarterly*, 101–114.

Schegloff, E. A. (1987b). Between micro and macro: Contexts and other connections. *The micro-macro link*, 207–234.

Schegloff, E. A. (1996). Issues of relevance for discourse analysis: Contingency in action, interaction and co-participant context. In *Computational and conversational discourse* (pp. 3–35). Springer.

Schegloff, E. A. (1997). Whose text? whose context? *Discourse & society*, *8*(2), 165–187.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289–327.

Schermerhorn, P., Scheutz, M., & Crowell, C. R. (2008). Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd acm/ieee international conference on human robot interaction* (pp. 263–270). ACM.

Scheutz, M., Cantrell, R., & Schermerhorn, P. (2011). Toward humanlike task-based dialogue processing for human robot interaction. *Ai Magazine*, *32*(4), 77–84.

Schlangen, D., & Skantze, G. (2011). A general, abstract model of incremental dialogue processing. *Special Issue on Incremental Processing in Dialogue*, *2*(1).

Scholtz, J. (2003). Theory and evaluation of human robot interactions. In *System sciences, 2003. proceedings of the 36th annual hawaii international conference on* (10–pp). IEEE.

Scholtz, J., Antonishek, B., & Young, J. D. (2005). Implementation of a situation awareness assessment tool for evaluation of human-robot interfaces. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *35*(4), 450–459.

Schroder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., … Pantic, M., et al. (2012). Building autonomous sensitive artificial listeners. *IEEE Transactions on Affective Computing*, *3*(2), 165–183.

Schröder, M., & Trouvain, J. (2003). The german text-to-speech synthesis system mary: A tool for research, development and teaching. *International Journal of Speech Technology*, *6*(4), 365–377.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., Pozzo, T., & Sandini, G. (2012). Measuring human-robot interaction through motor resonance. *International Journal of Social Robotics*, *4*(3), 223–234. doi:10.1007/s12369-012-0143-1

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., & Sandini, G. (2012). Anticipatory gaze in human-robot interactions. In *Gaze in hri from modeling to communication" workshop at the 7th acm/ieee international conference on human-robot interaction, boston, massachusetts, usa.*

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(3), 417–424.

Sekmen, A., & Challa, P. (2013). Assessment of adaptive human–robot interactions. *Knowledge-based systems*, *42*, 49–59.

Shiomi, M., Nakagawa, K., & Hagita, N. (2013). Design of a gaze behavior at a small mistake moment for a robot. *Interaction Studies*, *14*(3), 317–328.

Shukla, D., Erkent, Ö., & Piater, J. (2017). Supervised learning of gesture-action associations for human-robot collaboration. In *Automatic face & gesture recognition (fg 2017), 2017 12th ieee international conference on* (pp. 778–783). IEEE.

Sidner, C. L., Lee, C., Morency, L.-P., & Forlines, C. (2006). The effect of head-nod recognition in human-robot conversation. In *Proceedings of the 1st acm sigchi/sigart conference on human-robot interaction* (pp. 290–296). ACM.

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive robotics: The influence of robot gender on human behavior. In *Intelligent robots and systems, 2009. iros 2009. ieee/rsj international conference on* (pp. 2563–2568). IEEE.

Simmons, R., Makatchev, M., Kirby, R., Lee, M. K., Fanaswala, I., Browning, B., … Sakr, M. (2011). Believable robot characters. *AI Magazine*, *32*(4), 39–52.

Sirkin, D., Fischer, K., Jensen, L. C., & Ju, W. (2015). How effective an odd message can be. In *Conference on human computation & crowdsourcing.*

Sirkin, D., Fischer, K., Jensen, L. C., & Ju, W. (2016). Eliciting conversation in robot vehicle interactions. In *Proceedings of the association for the advancement of artifi-*

cial intelligence spring symposium series: Enabling computing research in socially intelligent human robot interaction (pp. 164–171).

Skantze, G., & Hjalmarsson, A. (2010). Towards incremental speech generation in dialogue systems. In *Proceedings of the 11th annual meeting of the special interest group on discourse and dialogue* (pp. 1–8). Association for Computational Linguistics.

Skantze, G., & Hjalmarsson, A. (2013). Towards incremental speech generation in conversational systems. *Computer Speech & Language*, *27*(1), 243–262.

Skantze, G., Hjalmarsson, A., & Oertel, C. (2014). Turn-taking, feedback and joint attention in situated human–robot interaction. *Speech Communication*, *65*, 50–66.

Skantze, G., Johansson, M., & Beskow, J. (2015). Exploring turn-taking cues in multi-party human-robot discussions about objects. In *Proceedings of the 2015 acm on international conference on multimodal interaction* (pp. 67–74). ACM.

Spexard, T. P., Hanheide, M., Li, S., & Wrede, B. (2008). Oops, something is wrong error detection and recovery for advanced human-robot-interaction. In *In proceedings of the icra workshop on social interaction with intelligent indoor robots (pasadena, ca).*

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, *93*(24), 13515–13522.

Stanton, C., & Stevens, C. J. (2014). Robot pressure: The impact of robot eye gaze and lifelike bodily movements upon decision-making and trust. In *International conference on social robotics* (pp. 330–339). Springer.

Streeck, J. (2015). Embodiment in human communication. *Annual Review of Anthropology*, *44*, 419–438.

Suarez, J., & Murphy, R. R. (2012). Hand gesture recognition with depth images: A review. In *Ro-man, 2012 ieee* (pp. 411–417). IEEE.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge university press.

Suchman, L. A. (2007). *Human-machine reconfigurations - plans and situated actions.* New York, NY: Cambridge University Press.

Swift-Spong, K., Wen, C. K. F., Spruijt-Metz, D., & Matarić, M. J. (2016). Comparing backstories of a socially assistive robot exercise buddy for adolescent youth. In *Robot and human interactive communication (ro-man), 2016 25th ieee international symposium on* (pp. 1013–1018). IEEE.

Szafir, D., & Mutlu, B. (2012). Pay attention!: Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 11–20). ACM.

Tanaka, F., Takahashi, T., Matsuzoe, S., Tazawa, N., & Morita, M. (2014). Telepresence robot helps children in communicating with teachers who speak a different language. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 399–406). ACM.

Trovato, G., Ham, J. R., Hashimoto, K., Ishii, H., & Takanishi, A. (2015). Investigating the effect of relative cultural distance on the acceptance of robots. In *International conference on social robotics* (pp. 664–673). Springer.

Tsai, V., Baumann, T., Pecune, F., & Casell, J. (2018). Faster responses are better responses: Introducing incrementality into sociable virtual personal assistants. In *Proceedings of the 2018 international workshop on spoken dialog system technology*.

Unhelkar, V. V., Siu, H. C., & Shah, J. A. (2014). Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction* (pp. 82–89). ACM.

Wang, E., Lignos, C., Vatsal, A., & Scassellati, B. (2006). Effects of head movement on perceptions of humanoid robot behavior. In *Proceedings of the 1st acm sigchi/sigart conference on human-robot interaction* (pp. 180–185). ACM.

Wang, L., Rau, P.-L. P., Evers, V., Robinson, B. K., & Hinds, P. (2010). When in rome: The role of culture & context in adherence to robot recommendations. In *Proceedings of the 5th acm/ieee international conference on human-robot interaction* (pp. 359–366). IEEE Press.

Weiss, A., Bernhaupt, R., Schwaiger, D., Altmaninger, M., Buchner, R., & Tscheligi, M. (2009). User experience evaluation with a wizard of oz approach: Technical and methodological considerations. In *Humanoid robots, 2009. humanoids 2009. 9th ieee-ras international conference on* (pp. 303–308). IEEE.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). Elan: A professional framework for multimodality research. In *Proceedings of lrec* (Vol. 2006, 5th).

Wrede, B., Buschkaemper, S., Muhl, C., & Rohlfing, K. J. (2006). Analyses of feedback in hri. *How People Talk to Computers, Robots, and Other Artificial Communication Partners*, 38.

Wu, E., Han, Y., Whitney, D., Oberlin, J., MacGlashan, J., & Tellex, S. (2015). Robotic social feedback for object specification. *AAAI: Artificial Intelligence for Human-Robot Interaction*.

Xu, Q., Li, L., & Wang, G. (2013). Designing engagement-aware agents for multiparty conversations. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 2233–2242). ACM.

Xu, T. L., Zhang, H., & Yu, C. (2016). See you see me: The role of eye contact in multimodal human-robot interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *6*(1), 2.

Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., & Kuzuoka, H. (2008). Precision timing in human-robot interaction: Coordination of head movement and utterance. Conference Paper. doi:10.1145/1357054.1357077

Yamazaki, K., Yamazaki, A., Okada, M., Kuno, Y., Kobayashi, Y., Hoshi, Y., … Heath, C. (2009). Revealing gauguin: Engaging visitors in robot guide's explanation in an art museum. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1437–1446). CHI '09. doi:10.1145/1518701.1518919

Yanco, H. A., & Drury, J. L. (2004). " where am i?" acquiring situation awareness using a remote robot platform. In *Ieee international conference on systems, man and cybernetics* (Vol. 3, pp. 2835–2840). IEEE.

Yu, C., Scheutz, M., & Schermerhorn, P. (2010). Investigating multimodal real-time patterns of joint attention in an hri word learning task. Conference Paper. IEEE Press.

Yu, Z., Bohus, D., & Horvitz, E. (2015). Incremental coordination: Attention-centric speech production in a physically situated conversational agent. In *Proceedings of the 16th annual meeting of the special interest group on discourse and dialogue* (pp. 402–406).

Zhang, J., Zheng, J., & Thalmann, N. M. (2018). Mcaem: Mixed-correlation analysis-based episodic memory for companion–user interactions. *The Visual Computer*, *34*(6-8), 1129–1141.

Zheng, K., Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2011). How many social robots can one operator control? In *Proceedings of the 6th international conference on human-robot interaction* (pp. 379–386). ACM.

Zheng, M., Moon, A., Croft, E. A., & Meng, M. Q.-H. (2015). Impacts of robot head gaze on robot-to-human handovers. *International Journal of Social Robotics*, *7*(5), 783–798. doi:10.1007/s12369-015-0305-z

zu Borgsen, S. M., Bernotat, J., & Wachsmuth, S. (2017). Hand in hand with robots: Differences between experienced and naive users in human-robot handover scenarios. In *International conference on social robotics* (pp. 587–596). Springer.

A. Regression Tables

A.1 Chapter 3

A.1.1 Subjective Ratings

Random condition as referent:

```
[1] "funBoring"     "LongShort"          "excitingTiring"          "IntEngaging"
[5] "intelligent"   "friendly"           "engaging"                "GameappropriateNotAppropriate"
[9] "GameFunBoring" "GameEasyDifficult"  "GameRepetitiveDynamic"   "GameEngagingBoring"

[[1]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min     1Q Median     3Q    Max
-2.399 -1.334 -0.140  1.080  2.446

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.4176     0.9360   0.446   0.6597
cnod          0.9410     0.8799   1.069   0.2960
cgaze         1.2915     0.7978   1.619   0.1191
GenderMale   -1.5364     0.7590  -2.024   0.0547 .
RoboXP        0.8449     0.3643   2.319   0.0296 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.69 on 23 degrees of freedom
Multiple R-squared:  0.2931,Adjusted R-squared:  0.1702
F-statistic: 2.384 on 4 and 23 DF,  p-value: 0.08093


[[2]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-2.5278 -0.1159  0.2683  0.5110  1.1589

Coefficients:
Estimate Std. Error t value   Pr(>|t|)
(Intercept)  3.54865    0.56255   6.308 0.00000296 ***
cnod        -0.71390    0.49734  -1.435      0.166
cgaze       -0.16713    0.46093  -0.363      0.721
```

```
GenderMale  -0.05165    0.42182  -0.122      0.904
RoboXP       0.14623    0.20740   0.705      0.489
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9325 on 21 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.09402,Adjusted R-squared:  -0.07855
F-statistic: 0.5448 on 4 and 21 DF,  p-value: 0.7047


[[3]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min     1Q Median     3Q    Max
-3.094 -1.240  0.228  1.160  2.139

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.26787    0.91813   2.470   0.0214 *
cnod         0.07580    0.86315   0.088   0.9308
cgaze       -0.05005    0.78258  -0.064   0.9496
GenderMale  -0.82093    0.74457  -1.103   0.2816
RoboXP       0.64279    0.35734   1.799   0.0852 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.658 on 23 degrees of freedom
Multiple R-squared:  0.1784,Adjusted R-squared:  0.03551
F-statistic: 1.249 on 4 and 23 DF,  p-value: 0.3185


[[4]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.7909 -0.7805  0.1038  0.8722  1.6187

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.44542    0.62481   3.914 0.000697 ***
cnod         0.27044    0.58740   0.460 0.649556
cgaze        0.37755    0.53257   0.709 0.485497
GenderMale   0.42782    0.50670   0.844 0.407190
RoboXP      -0.03207    0.24318  -0.132 0.896221
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.128 on 23 degrees of freedom
Multiple R-squared:  0.06288,Adjusted R-squared:  -0.1001
F-statistic: 0.3859 on 4 and 23 DF,  p-value: 0.8165
```

```
[[5]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8762 -0.7614  0.1525  1.2049  1.6378

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.73302    0.64723   4.223 0.000323 ***
cnod        -0.09824    0.60848  -0.161 0.873143
cgaze       -0.38799    0.55168  -0.703 0.488938
GenderMale   0.09167    0.52488   0.175 0.862887
RoboXP       0.01716    0.25191   0.068 0.946266
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.169 on 23 degrees of freedom
Multiple R-squared:  0.03097,Adjusted R-squared:  -0.1376
F-statistic: 0.1838 on 4 and 23 DF,  p-value: 0.9445


[[6]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8230 -0.7116 -0.1390  0.6588  1.5774

Coefficients:
Estimate Std. Error t value   Pr(>|t|)
(Intercept)   3.2233     0.5228   6.165 0.00000332 ***
cnod         -0.1412     0.4989  -0.283     0.7798
cgaze        -1.0802     0.4440  -2.433     0.0235 *
GenderMale    0.6583     0.4253   1.548     0.1360
RoboXP       -0.4004     0.2031  -1.971     0.0615 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9404 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.2704,Adjusted R-squared:  0.1377
F-statistic: 2.038 on 4 and 22 DF,  p-value: 0.124


[[7]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-1.90698 -0.88598 -0.00818  0.81311  1.43162
```

```
Coefficients:
Estimate Std. Error t value  Pr(>|t|)
(Intercept)  3.21758    0.58211   5.527 0.0000127 ***
cnod         0.08737    0.54726   0.160    0.8746
cgaze       -0.71742    0.49617  -1.446    0.1617
GenderMale   0.98930    0.47207   2.096    0.0473 *
RoboXP      -0.31059    0.22656  -1.371    0.1836
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.051 on 23 degrees of freedom
Multiple R-squared:  0.2162,Adjusted R-squared:  0.07984
F-statistic: 1.586 on 4 and 23 DF,  p-value: 0.2118


[[8]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.9984 -0.5455  0.2319  0.6587  1.3851

Coefficients:
Estimate Std. Error t value     Pr(>|t|)
(Intercept)  4.31736    0.55329   7.803 0.0000000656 ***
cnod         0.42835    0.52016   0.824        0.419
cgaze        0.27707    0.47160   0.588        0.563
GenderMale  -0.04482    0.44870  -0.100        0.921
RoboXP      -0.17561    0.21534  -0.816        0.423
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9992 on 23 degrees of freedom
Multiple R-squared:  0.06948,Adjusted R-squared:  -0.09235
F-statistic: 0.4293 on 4 and 23 DF,  p-value: 0.7859


[[9]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.5525 -0.4698 -0.2891  0.3083  2.1458

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.4072     0.5502   4.375 0.000221 ***
cnod         -0.2280     0.5173  -0.441 0.663500
cgaze        -0.6983     0.4690  -1.489 0.150099
GenderMale   -0.4713     0.4462  -1.056 0.301852
RoboXP        0.1453     0.2142   0.678 0.504297
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9937 on 23 degrees of freedom
```

```
Multiple R-squared:  0.1665,Adjusted R-squared:  0.02159
F-statistic: 1.149 on 4 and 23 DF,  p-value: 0.3588


[[10]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q Median     3Q    Max
-2.285 -1.047  0.184  0.904  2.149

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.9131     0.6948   2.754   0.0113 *
cnod         -0.3126     0.6532  -0.479   0.6367
cgaze         0.5038     0.5922   0.851   0.4037
GenderMale   -0.7044     0.5634  -1.250   0.2238
RoboXP        0.4340     0.2704   1.605   0.1222
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.255 on 23 degrees of freedom
Multiple R-squared:  0.1222,Adjusted R-squared:  -0.03048
F-statistic: 0.8004 on 4 and 23 DF,  p-value: 0.5374


[[11]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-2.39896 -0.31892 -0.06428  0.49953  2.14167

Coefficients:
Estimate Std. Error t value   Pr(>|t|)
(Intercept)   3.2951     0.5554   5.933 0.00000477 ***
cnod         -0.4498     0.5222  -0.862     0.3979
cgaze        -0.4273     0.4734  -0.903     0.3761
GenderMale   -1.1133     0.4504  -2.472     0.0213 *
RoboXP        0.1039     0.2162   0.481     0.6354
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.003 on 23 degrees of freedom
Multiple R-squared:  0.2793,Adjusted R-squared:  0.1539
F-statistic: 2.228 on 4 and 23 DF,  p-value: 0.09744


[[12]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
```

```
Min      1Q  Median      3Q     Max
-1.3575 -0.6949 -0.2199  0.5884  1.8883

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.2492     0.5396   4.169   0.0004 ***
cnod          0.3400     0.5002   0.680   0.5038
cgaze        -0.2458     0.4584  -0.536   0.5972
GenderMale   -0.3273     0.4692  -0.698   0.4928
RoboXP        0.1082     0.2165   0.500   0.6221
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9608 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.09809,Adjusted R-squared:  -0.06589
F-statistic: 0.5982 on 4 and 22 DF,  p-value: 0.6678
```

Contingent gaze condition as referent:

```
[[1]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min     1Q Median     3Q    Max
-2.399 -1.334 -0.140  1.080  2.446

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.7090     0.7507   2.277   0.0324 *
cnod         -0.3504     0.9925  -0.353   0.7273
random       -1.2915     0.7978  -1.619   0.1191
GenderMale   -1.5364     0.7590  -2.024   0.0547 .
RoboXP        0.8449     0.3643   2.319   0.0296 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.69 on 23 degrees of freedom
Multiple R-squared:  0.2931,Adjusted R-squared:  0.1702
F-statistic: 2.384 on 4 and 23 DF,  p-value: 0.08093


[[2]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-2.5278 -0.1159  0.2683  0.5110  1.1589

Coefficients:
Estimate Std. Error t value     Pr(>|t|)
(Intercept)  3.38152    0.42371   7.981 0.0000000856 ***
cnod        -0.54677    0.56082  -0.975        0.341
random       0.16713    0.46093   0.363        0.721
GenderMale  -0.05165    0.42182  -0.122        0.904
```

```
RoboXP       0.14623    0.20740   0.705        0.489
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9325 on 21 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.09402,Adjusted R-squared:  -0.07855
F-statistic: 0.5448 on 4 and 21 DF,  p-value: 0.7047


[[3]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min     1Q Median     3Q    Max
-3.094 -1.240  0.228  1.160  2.139

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.21782    0.73638   3.012  0.00622 **
cnod         0.12585    0.97360   0.129  0.89827
random       0.05005    0.78258   0.064  0.94956
GenderMale  -0.82093    0.74457  -1.103  0.28163
RoboXP       0.64279    0.35734   1.799  0.08519 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.658 on 23 degrees of freedom
Multiple R-squared:  0.1784,Adjusted R-squared:  0.03551
F-statistic: 1.249 on 4 and 23 DF,  p-value: 0.3185


[[4]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.7909 -0.7805  0.1038  0.8722  1.6187

Coefficients:
Estimate Std. Error t value   Pr(>|t|)
(Intercept)  2.82297    0.50113   5.633 0.00000984 ***
cnod        -0.10711    0.66256  -0.162      0.873
random      -0.37755    0.53257  -0.709      0.485
GenderMale   0.42782    0.50670   0.844      0.407
RoboXP      -0.03207    0.24318  -0.132      0.896
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.128 on 23 degrees of freedom
Multiple R-squared:  0.06288,Adjusted R-squared:  -0.1001
F-statistic: 0.3859 on 4 and 23 DF,  p-value: 0.8165


[[5]]
```

```
Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8762 -0.7614  0.1525  1.2049  1.6378

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.34504    0.51911   4.517 0.000155 ***
cnod         0.28974    0.68634   0.422 0.676831
random       0.38799    0.55168   0.703 0.488938
GenderMale   0.09167    0.52488   0.175 0.862887
RoboXP       0.01716    0.25191   0.068 0.946266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.169 on 23 degrees of freedom
Multiple R-squared:  0.03097,Adjusted R-squared:  -0.1376
F-statistic: 0.1838 on 4 and 23 DF,  p-value: 0.9445


[[6]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8230 -0.7116 -0.1390  0.6588  1.5774

Coefficients:
Estimate Std. Error t value  Pr(>|t|)
(Intercept)   2.1431     0.4193   5.111 0.0000403 ***
cnod          0.9391     0.5590   1.680    0.1071
random        1.0802     0.4440   2.433    0.0235 *
GenderMale    0.6583     0.4253   1.548    0.1360
RoboXP       -0.4004     0.2031  -1.971    0.0615 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9404 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.2704,Adjusted R-squared:  0.1377
F-statistic: 2.038 on 4 and 22 DF,  p-value: 0.124


[[7]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-1.90698 -0.88598 -0.00818  0.81311  1.43162

Coefficients:
```

```
Estimate Std. Error t value  Pr(>|t|)
(Intercept)   2.5002     0.4669   5.355 0.0000194 ***
cnod          0.8048     0.6173   1.304    0.2052
random        0.7174     0.4962   1.446    0.1617
GenderMale    0.9893     0.4721   2.096    0.0473 *
RoboXP       -0.3106     0.2266  -1.371    0.1836
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.051 on 23 degrees of freedom
Multiple R-squared:  0.2162,Adjusted R-squared:  0.07984
F-statistic: 1.586 on 4 and 23 DF,  p-value: 0.2118


[[8]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.9984 -0.5455  0.2319  0.6587  1.3851

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)  4.59443    0.44376  10.353 0.000000000395 ***
cnod         0.15128    0.58672   0.258          0.799
random      -0.27707    0.47160  -0.588          0.563
GenderMale  -0.04482    0.44870  -0.100          0.921
RoboXP      -0.17561    0.21534  -0.816          0.423
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9992 on 23 degrees of freedom
Multiple R-squared:  0.06948,Adjusted R-squared:  -0.09235
F-statistic: 0.4293 on 4 and 23 DF,  p-value: 0.7859


[[9]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.5525 -0.4698 -0.2891  0.3083  2.1458

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.7089     0.4413   3.872 0.000772 ***
cnod          0.4703     0.5835   0.806 0.428493
random        0.6983     0.4690   1.489 0.150099
GenderMale   -0.4713     0.4462  -1.056 0.301852
RoboXP        0.1453     0.2142   0.678 0.504297
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9937 on 23 degrees of freedom
Multiple R-squared:  0.1665,Adjusted R-squared:  0.02159
```

```
F-statistic: 1.149 on 4 and 23 DF,  p-value: 0.3588


[[10]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min     1Q Median     3Q    Max
-2.285 -1.047  0.184  0.904  2.149

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.4169     0.5572   4.337 0.000243 ***
cnod         -0.8164     0.7367  -1.108 0.279270
random       -0.5038     0.5922  -0.851 0.403695
GenderMale   -0.7044     0.5634  -1.250 0.223785
RoboXP        0.4340     0.2704   1.605 0.122152
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.255 on 23 degrees of freedom
Multiple R-squared:  0.1222,Adjusted R-squared:  -0.03048
F-statistic: 0.8004 on 4 and 23 DF,  p-value: 0.5374


[[11]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-2.39896 -0.31892 -0.06428  0.49953  2.14167

Coefficients:
Estimate Std. Error t value   Pr(>|t|)
(Intercept)  2.86779    0.44546   6.438 0.00000144 ***
cnod        -0.02255    0.58896  -0.038     0.9698
random       0.42728    0.47341   0.903     0.3761
GenderMale  -1.11334    0.45042  -2.472     0.0213 *
RoboXP       0.10388    0.21617   0.481     0.6354
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.003 on 23 degrees of freedom
Multiple R-squared:  0.2793,Adjusted R-squared:  0.1539
F-statistic: 2.228 on 4 and 23 DF,  p-value: 0.09744


[[12]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
```

```
-1.3575 -0.6949 -0.2199  0.5884  1.8883

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.0034     0.4546   4.407 0.000223 ***
cnod          0.5858     0.5672   1.033 0.312943
random        0.2458     0.4584   0.536 0.597177
GenderMale   -0.3273     0.4692  -0.698 0.492768
RoboXP        0.1082     0.2165   0.500 0.622057
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9608 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.09809,Adjusted R-squared:  -0.06589
F-statistic: 0.5982 on 4 and 22 DF,  p-value: 0.6678
```

A.1.2 Linguistic Analysis

Random Gaze as referent:

```
[1] "utTurn" "expoUt" "mlu"

[[1]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.1039 -0.7434 -0.1123  0.3064  1.7602

Coefficients:
Estimate Std. Error t value    Pr(>|t|)
(Intercept)  3.43073    0.48961   7.007 0.000000385 ***
cnod         0.43695    0.46029   0.949      0.3523
cgaze       -0.85745    0.41732  -2.055      0.0514 .
GenderMale  -0.08296    0.39706  -0.209      0.8363
RoboXP      -0.19095    0.19056  -1.002      0.3267
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8842 on 23 degrees of freedom
Multiple R-squared:  0.2384,Adjusted R-squared:  0.106
F-statistic:   1.8 on 4 and 23 DF,  p-value: 0.1632


[[2]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-15.5071  -3.0894  -0.8966   2.8456  20.1233

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.213      4.101   0.540  0.59455
```

```
cnod           2.539      3.855   0.659  0.51664
cgaze         11.424      3.495   3.268  0.00338 **
GenderMale    -9.342      3.326  -2.809  0.00996 **
RoboXP         1.870      1.596   1.171  0.25343
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.406 on 23 degrees of freedom
Multiple R-squared:  0.4462,Adjusted R-squared:  0.3499
F-statistic: 4.632 on 4 and 23 DF,  p-value: 0.006859


[[3]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-2.3488 -0.8180 -0.1178  0.4909  4.7227

Coefficients:
Estimate Std. Error t value  Pr(>|t|)
(Intercept)   5.0233     0.9038   5.558 0.0000118 ***
cnod         -0.7948     0.8497  -0.935    0.3593
cgaze         0.9569     0.7704   1.242    0.2267
GenderMale   -0.8314     0.7330  -1.134    0.2683
RoboXP        0.6588     0.3518   1.873    0.0738 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.632 on 23 degrees of freedom
Multiple R-squared:  0.1753,Adjusted R-squared:  0.03183
F-statistic: 1.222 on 4 and 23 DF,  p-value: 0.3288

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Contingent gaze as referent:

```
inter.m <- lapply(interList, function(x) {
lm(substitute(i~Gender+RoboXP+cnod+random, list(i = as.name(x))), data = d2)})
lapply(inter.m, summary)

[[1]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.1039 -0.7434 -0.1123  0.3064  1.7602

Coefficients:
Estimate Std. Error t value  Pr(>|t|)
(Intercept)  2.57328    0.39268   6.553 0.0000011 ***
cnod         1.29440    0.51919   2.493    0.0203 *
random       0.85745    0.41732   2.055    0.0514 .
GenderMale  -0.08296    0.39706  -0.209    0.8363
RoboXP      -0.19095    0.19056  -1.002    0.3267
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8842 on 23 degrees of freedom
Multiple R-squared:  0.2384,Adjusted R-squared:  0.106
F-statistic:   1.8 on 4 and 23 DF,  p-value: 0.1632


[[2]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-15.5071  -3.0894  -0.8966   2.8456  20.1233

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   13.637      3.289   4.146 0.000391 ***
cnod          -8.885      4.349  -2.043 0.052662 .
random       -11.424      3.495  -3.268 0.003378 **
GenderMale    -9.342      3.326  -2.809 0.009957 **
RoboXP         1.870      1.596   1.171 0.253434
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.406 on 23 degrees of freedom
Multiple R-squared:  0.4462,Adjusted R-squared:  0.3499
F-statistic: 4.632 on 4 and 23 DF,  p-value: 0.006859


[[3]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-2.3488 -0.8180 -0.1178  0.4909  4.7227

Coefficients:
Estimate Std. Error t value     Pr(>|t|)
(Intercept)   5.9802     0.7249   8.250 0.0000000252 ***
cnod         -1.7516     0.9584  -1.828       0.0806 .
random       -0.9569     0.7704  -1.242       0.2267
GenderMale   -0.8314     0.7330  -1.134       0.2683
RoboXP        0.6588     0.3518   1.873       0.0738 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.632 on 23 degrees of freedom
Multiple R-squared:  0.1753,Adjusted R-squared:  0.03183
F-statistic: 1.222 on 4 and 23 DF,  p-value: 0.3288
```

A.1.3 Interpersonal I

Random Gaze as referent:

```
[1] "questUt" "tagUt"   "tagUt"

[[1]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + random,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-10.119  -2.964  -1.173   2.041  13.531

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    9.165      2.435   3.764  0.00101 **
GenderMale     4.041      2.462   1.641  0.11432
RoboXP        -3.777      1.182  -3.196  0.00401 **
cnod           5.270      3.219   1.637  0.11528
random         4.731      2.588   1.828  0.08050 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.483 on 23 degrees of freedom
Multiple R-squared:  0.3293,Adjusted R-squared:  0.2126
F-statistic: 2.823 on 4 and 23 DF,  p-value: 0.04847


[[2]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + random,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8672 -0.8836 -0.3220  0.9471  3.0984

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.2738     0.5913   3.845 0.000825 ***
GenderMale   -0.4899     0.5979  -0.819 0.420958
RoboXP       -0.5277     0.2869  -1.839 0.078865 .
cnod          0.6488     0.7818   0.830 0.415108
random        0.7111     0.6284   1.132 0.269473
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.331 on 23 degrees of freedom
Multiple R-squared:  0.2606,Adjusted R-squared:  0.132
F-statistic: 2.027 on 4 and 23 DF,  p-value: 0.1241


[[3]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + random,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8672 -0.8836 -0.3220  0.9471  3.0984
```

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.2738     0.5913   3.845 0.000825 ***
GenderMale   -0.4899     0.5979  -0.819 0.420958
RoboXP       -0.5277     0.2869  -1.839 0.078865 .
cnod          0.6488     0.7818   0.830 0.415108
random        0.7111     0.6284   1.132 0.269473
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.331 on 23 degrees of freedom
Multiple R-squared:  0.2606,Adjusted R-squared:  0.132
F-statistic: 2.027 on 4 and 23 DF,  p-value: 0.1241
```

Contingent gaze as referent:

```
> inter.m <- lapply(interList, function(x) {
+     lm(substitute(i~Gender+RoboXP+cnod+cgaze, list(i = as.name(x))), data = d2)})
> lapply(inter.m, summary)
[[1]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + cgaze, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-10.119  -2.964  -1.173   2.041  13.531

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.8961     3.0360   4.577 0.000134 ***
GenderMale    4.0413     2.4621   1.641 0.114319
RoboXP       -3.7771     1.1816  -3.196 0.004013 **
cnod          0.5384     2.8542   0.189 0.852047
cgaze        -4.7313     2.5878  -1.828 0.080503 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.483 on 23 degrees of freedom
Multiple R-squared:  0.3293,Adjusted R-squared:  0.2126
F-statistic: 2.823 on 4 and 23 DF,  p-value: 0.04847


[[2]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + cgaze, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8672 -0.8836 -0.3220  0.9471  3.0984

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.98484    0.73725   4.049 0.000499 ***
GenderMale  -0.48992    0.59788  -0.819 0.420958
RoboXP      -0.52768    0.28694  -1.839 0.078865 .
cnod        -0.06225    0.69310  -0.090 0.929213
```

```
cgaze       -0.71109    0.62840  -1.132 0.269473
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.331 on 23 degrees of freedom
Multiple R-squared:  0.2606,Adjusted R-squared:  0.132
F-statistic: 2.027 on 4 and 23 DF,  p-value: 0.1241


[[3]]

Call:
lm(formula = substitute(i ~ Gender + RoboXP + cnod + cgaze, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8672 -0.8836 -0.3220  0.9471  3.0984

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.98484    0.73725   4.049 0.000499 ***
GenderMale  -0.48992    0.59788  -0.819 0.420958
RoboXP      -0.52768    0.28694  -1.839 0.078865 .
cnod        -0.06225    0.69310  -0.090 0.929213
cgaze       -0.71109    0.62840  -1.132 0.269473
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.331 on 23 degrees of freedom
Multiple R-squared:  0.2606,Adjusted R-squared:  0.132
F-statistic: 2.027 on 4 and 23 DF,  p-value: 0.1241
```

A.1.4 Interpersonal II

Random Gaze as referent:

```
> proModel <- lm(firstPerProSingUt~cnod+cgaze+Gender+RoboXP, data=d2)
> summary(proModel)

Call:
lm(formula = firstPerProSingUt ~ cnod + cgaze + Gender + RoboXP,
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-14.296  -5.468  -1.724   4.003  23.314

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  16.1363     5.6048   2.879  0.00847 **
cnod        -16.8622     5.2692  -3.200  0.00398 **
cgaze        -8.9154     4.7773  -1.866  0.07482 .
GenderMale    0.3979     4.5453   0.088  0.93101
RoboXP        3.0373     2.1814   1.392  0.17713
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.12 on 23 degrees of freedom
Multiple R-squared:  0.3426,Adjusted R-squared:  0.2282
F-statistic: 2.996 on 4 and 23 DF,  p-value: 0.03972
```

```
> proModel2 <- lm(firstPlural~cnod+cgaze+Gender+RoboXP, data=d2)
> summary(proModel2)

Call:
lm(formula = firstPlural ~ cnod + cgaze + Gender + RoboXP, data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-11.921  -7.527  -0.831   4.884  39.143

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  15.0619     6.3500   2.372   0.0269 *
cnod          2.0768     5.9562   0.349   0.7306
cgaze        -1.8277     5.4996  -0.332   0.7428
GenderMale   -0.9275     5.1477  -0.180   0.8587
RoboXP       -1.3135     2.4689  -0.532   0.6001
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.42 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.02438,Adjusted R-squared:  -0.153
F-statistic: 0.1375 on 4 and 22 DF,  p-value: 0.9666

> proModel3 <- lm(youut~cnod+cgaze+Gender+RoboXP, data=d2)
> summary(proModel3)

Call:
lm(formula = youut ~ cnod + cgaze + Gender + RoboXP, data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-12.5002  -6.2173  -0.8689   4.5771  31.1428

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  16.3662     6.0534   2.704   0.0127 *
cnod         -2.3460     5.6909  -0.412   0.6840
cgaze         0.3002     5.1597   0.058   0.9541
GenderMale   -0.0603     4.9091  -0.012   0.9903
RoboXP       -1.1216     2.3560  -0.476   0.6385
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.93 on 23 degrees of freedom
Multiple R-squared:  0.04732,Adjusted R-squared:  -0.1184
F-statistic: 0.2856 on 4 and 23 DF,  p-value: 0.8843
```

Contingent Gaze as referent:

```
> proModel <- lm(firstPerProSingUt~cnod+random+Gender+RoboXP, data=d2)
> summary(proModel)

Call:
lm(formula = firstPerProSingUt ~ cnod + random + Gender + RoboXP,
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-14.296  -5.468  -1.724   4.003  23.314
```

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.2209     4.4952   1.606   0.1218
cnod         -7.9468     5.9434  -1.337   0.1943
random        8.9154     4.7773   1.866   0.0748 .
GenderMale    0.3979     4.5453   0.088   0.9310
RoboXP        3.0373     2.1814   1.392   0.1771
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.12 on 23 degrees of freedom
Multiple R-squared:  0.3426,Adjusted R-squared:  0.2282
F-statistic: 2.996 on 4 and 23 DF,  p-value: 0.03972

> proModel2 <- lm(firstPlural~cnod+random+Gender+RoboXP, data=d2)
> summary(proModel2)

Call:
lm(formula = firstPlural ~ cnod + random + Gender + RoboXP, data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-11.921  -7.527  -0.831   4.884  39.143

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.2342     5.3340   2.481   0.0212 *
cnod          3.9045     6.7483   0.579   0.5687
random        1.8277     5.4996   0.332   0.7428
GenderMale   -0.9275     5.1477  -0.180   0.8587
RoboXP       -1.3135     2.4689  -0.532   0.6001
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.42 on 22 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.02438,Adjusted R-squared:  -0.153
F-statistic: 0.1375 on 4 and 22 DF,  p-value: 0.9666

> proModel3 <- lm(youut~cnod+random+Gender+RoboXP, data=d2)
> summary(proModel3)

Call:
lm(formula = youut ~ cnod + random + Gender + RoboXP, data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-12.5002  -6.2173  -0.8689   4.5771  31.1428

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  16.6664     4.8551   3.433  0.00227 **
cnod         -2.6462     6.4191  -0.412  0.68398
random       -0.3002     5.1597  -0.058  0.95411
GenderMale   -0.0603     4.9091  -0.012  0.99030
RoboXP       -1.1216     2.3560  -0.476  0.63853
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.93 on 23 degrees of freedom
Multiple R-squared:  0.04732,Adjusted R-squared:  -0.1184
```

```
F-statistic: 0.2856 on 4 and 23 DF,  p-value: 0.8843
```

A.1.5 Confusion

Random gaze as referent:

```
[1] "toexpUT"     "firstPlural" "srepUt"

> confList <- names(d2)[c(69,86,87)]
> conf.m <- lapply(confList, function(x) {
+     lm(substitute(i~cnod+cgaze+Gender+RoboXP, list(i = as.name(x))), data = d2)})
> lapply(conf.m, summary)
[[1]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-22.653  -6.816  -2.025   7.987  20.640

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.2783     6.9408   2.489   0.0205 *
cnod        -16.3113     6.5252  -2.500   0.0200 *
cgaze        -0.2348     5.9161  -0.040   0.9687
GenderMale    7.6228     5.6287   1.354   0.1888
RoboXP        3.0332     2.7014   1.123   0.2731
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 12.53 on 23 degrees of freedom
Multiple R-squared:  0.2761,Adjusted R-squared:  0.1503
F-statistic: 2.194 on 4 and 23 DF,  p-value: 0.1016


[[2]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-11.1015  -3.1030  -0.5464   2.4951  14.7074

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.142      3.396   0.925   0.3644
cnod           5.643      3.192   1.768   0.0904 .
cgaze          1.675      2.894   0.579   0.5684
GenderMale    -5.614      2.754  -2.039   0.0531 .
RoboXP         1.158      1.322   0.876   0.3900
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.133 on 23 degrees of freedom
Multiple R-squared:  0.2462,Adjusted R-squared:  0.1152
F-statistic: 1.879 on 4 and 23 DF,  p-value: 0.1484
```

```
[[3]]

Call:
lm(formula = substitute(i ~ cnod + cgaze + Gender + RoboXP, list(i = as.name(x))),
data = d2)

Residuals:
Min     1Q Median     3Q    Max
-4.723 -2.833 -0.727  1.599 10.065

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.7394     2.3081   3.353  0.00275 **
cnod          1.2876     2.1699   0.593  0.55872
cgaze        -1.1541     1.9673  -0.587  0.56318
GenderMale    5.8200     1.8718   3.109  0.00494 **
RoboXP       -3.0167     0.8983  -3.358  0.00272 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.168 on 23 degrees of freedom
Multiple R-squared:  0.3956,Adjusted R-squared:  0.2905
F-statistic: 3.764 on 4 and 23 DF,  p-value: 0.01696
```

Contingent Gaze as referent:

```
> conf.m <- lapply(confList, function(x) {
+     lm(substitute(i~cnod+random+Gender+RoboXP, list(i = as.name(x))), data = d2)})
> lapply(conf.m, summary)
[[1]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-22.653  -6.816  -2.025   7.987  20.640

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.0434     5.5668   3.062  0.00553 **
cnod        -16.0764     7.3601  -2.184  0.03940 *
random        0.2348     5.9161   0.040  0.96868
GenderMale    7.6228     5.6287   1.354  0.18881
RoboXP        3.0332     2.7014   1.123  0.27309
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12.53 on 23 degrees of freedom
Multiple R-squared:  0.2761,Adjusted R-squared:  0.1503
F-statistic: 2.194 on 4 and 23 DF,  p-value: 0.1016


[[2]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min       1Q   Median       3Q      Max
```

```
-11.1015  -3.1030  -0.5464   2.4951  14.7074

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.817      2.724   1.769   0.0902 .
cnod           3.968      3.601   1.102   0.2819
random        -1.675      2.894  -0.579   0.5684
GenderMale    -5.614      2.754  -2.039   0.0531 .
RoboXP         1.158      1.322   0.876   0.3900
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.133 on 23 degrees of freedom
Multiple R-squared:  0.2462,Adjusted R-squared:  0.1152
F-statistic: 1.879 on 4 and 23 DF,  p-value: 0.1484


[[3]]

Call:
lm(formula = substitute(i ~ cnod + random + Gender + RoboXP,
list(i = as.name(x))), data = d2)

Residuals:
Min     1Q Median     3Q    Max
-4.723 -2.833 -0.727  1.599 10.065

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.5854     1.8512   3.557  0.00168 **
cnod          2.4416     2.4475   0.998  0.32886
random        1.1541     1.9673   0.587  0.56318
GenderMale    5.8200     1.8718   3.109  0.00494 **
RoboXP       -3.0167     0.8983  -3.358  0.00272 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.168 on 23 degrees of freedom
Multiple R-squared:  0.3956,Adjusted R-squared:  0.2905
F-statistic: 3.764 on 4 and 23 DF,  p-value: 0.01696
```

A.1.6 Gaze

Random condition as referent:

```
> gazeN <- lm(gazeNut~cgaze+cnod, data=d2)
> summary(gazeN)

Call:
lm(formula = gazeNut ~ cgaze + cnod, data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-4.5122 -1.4266  0.1098  1.1661  4.2951

Coefficients:
Estimate Std. Error t value    Pr(>|t|)
(Intercept)  4.660757   0.689314   6.761 0.000000437 ***
cgaze       -0.003243   0.974837  -0.003       0.997
cnod        -2.193648   1.033971  -2.122       0.044 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Residual standard error: 2.18 on 25 degrees of freedom
Multiple R-squared:  0.1877,Adjusted R-squared:  0.1228
F-statistic: 2.889 on 2 and 25 DF,  p-value: 0.07433

> gazeDuravg <- lm(gazeToRobInstrAvgDur~cgaze+cnod, data=d2)
> summary(gazeDuravg)

Call:
lm(formula = gazeToRobInstrAvgDur ~ cgaze + cnod, data = d2)

Residuals:
Min     1Q Median     3Q    Max
-2.785 -0.813 -0.125  0.404  4.426

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.6340     0.4908   3.329   0.0027 **
cgaze        -0.1280     0.6941  -0.184   0.8552
cnod          1.6010     0.7362   2.175   0.0393 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.552 on 25 degrees of freedom
Multiple R-squared:  0.2091,Adjusted R-squared:  0.1458
F-statistic: 3.305 on 2 and 25 DF,  p-value: 0.05326

> gazedur <- lm(gazeDur~cgaze+cnod, data=d2)
> summary(gazedur)

Call:
lm(formula = gazeDur ~ cgaze + cnod, data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-15.3277  -6.7296  -0.4192   3.2035  26.4988

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.6281     3.4702   3.639  0.00124 **
cgaze        -0.5407     4.9077  -0.110  0.91315
cnod          3.3068     5.2054   0.635  0.53102
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.97 on 25 degrees of freedom
Multiple R-squared:  0.02417,Adjusted R-squared:  -0.05389
F-statistic: 0.3097 on 2 and 25 DF,  p-value: 0.7365
```

Contingent gaze condition as referent:

```
> gazeN <- lm(gazeNut~random+cnod, data=d2)
> summary(gazeN)

Call:
lm(formula = gazeNut ~ random + cnod, data = d2)

Residuals:
Min      1Q  Median      3Q     Max
-4.5122 -1.4266  0.1098  1.1661  4.2951

Coefficients:
```

```
Estimate Std. Error t value    Pr(>|t|)
(Intercept)  4.657513   0.689314   6.757 0.000000442 ***
random       0.003243   0.974837   0.003      0.9974
cnod        -2.190405   1.033971  -2.118      0.0443 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.18 on 25 degrees of freedom
Multiple R-squared:  0.1877,Adjusted R-squared:  0.1228
F-statistic: 2.889 on 2 and 25 DF,  p-value: 0.07433

> gazeDuravg <- lm(gazeToRobInstrAvgDur~random+cnod, data=d2)
> summary(gazeDuravg)

Call:
lm(formula = gazeToRobInstrAvgDur ~ random + cnod, data = d2)

Residuals:
Min     1Q Median     3Q    Max
-2.785 -0.813 -0.125  0.404  4.426

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.5060     0.4908   3.068  0.00512 **
random         0.1280     0.6941   0.184  0.85518
cnod           1.7290     0.7362   2.349  0.02706 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.552 on 25 degrees of freedom
Multiple R-squared:  0.2091,Adjusted R-squared:  0.1458
F-statistic: 3.305 on 2 and 25 DF,  p-value: 0.05326

> gazedur <- lm(gazeDur~random+cnod, data=d2)
> summary(gazedur)

Call:
lm(formula = gazeDur ~ random + cnod, data = d2)

Residuals:
Min       1Q   Median       3Q      Max
-15.3277  -6.7296  -0.4192   3.2035  26.4988

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.0874     3.4702   3.483  0.00184 **
random        0.5407     4.9077   0.110  0.91315
cnod          3.8476     5.2054   0.739  0.46670
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.97 on 25 degrees of freedom
Multiple R-squared:  0.02417,Adjusted R-squared:  -0.05389
F-statistic: 0.3097 on 2 and 25 DF,  p-value: 0.7365
```

A.2 Chapter 4

```
 [1] "robot_aware" "Post_soc2"   "Post_inter2" "Post_comp2"   "Post_reli2"   "Post_know2"   "Post_pintel" "competence"   "Pos

> modQuest <- lapply(varList3, function(x) {
```

```
+     lm(substitute(i~conAware+gender+prev_exp, list(i = as.name(x))), data = d1)})
> lapply(modQuest, summary)
[[1]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min       1Q   Median       3Q      Max
-1.51750 -0.08025  0.12394  0.31505  0.51924

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)  4.68495    0.16115  29.072 <0.0000000000000002 ***
conAware     0.35856    0.13930   2.574              0.0132 *
genderm     -0.20419    0.16090  -1.269              0.2105
prev_expyes  0.03674    0.14480   0.254              0.8008
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4879 on 48 degrees of freedom
Multiple R-squared:  0.1688,Adjusted R-squared:  0.1169
F-statistic:  3.25 on 3 and 48 DF,  p-value: 0.02973


[[2]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-3.8938 -0.7497  0.1990  1.1062  2.9850

Coefficients:
Estimate Std. Error t value        Pr(>|t|)
(Intercept)   4.5871     0.4967   9.236 0.0000000000032 ***
conAware      0.8788     0.4293   2.047          0.0462 *
genderm      -0.7464     0.4959  -1.505          0.1388
prev_expyes   0.1743     0.4463   0.391          0.6978
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.504 on 48 degrees of freedom
Multiple R-squared:  0.1405,Adjusted R-squared:  0.0868
F-statistic: 2.616 on 3 and 48 DF,  p-value: 0.06175


[[3]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-3.4775 -0.5109  0.3665  0.7373  1.5225

Coefficients:
```

```
Estimate Std. Error t value           Pr(>|t|)
(Intercept)   5.6112     0.3462  16.209 <0.0000000000000002 ***
conAware      0.6293     0.2992   2.103             0.0407 *
genderm      -0.1337     0.3456  -0.387             0.7005
prev_expyes   0.1560     0.3110   0.502             0.6182
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.048 on 48 degrees of freedom
Multiple R-squared:  0.1035,Adjusted R-squared:  0.04748
F-statistic: 1.847 on 3 and 48 DF,  p-value: 0.1512


[[4]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-3.6213 -0.6082  0.3787  1.0817  1.5499

Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)  5.80649    0.41834  13.880 <0.0000000000000002 ***
conAware     0.11181    0.36162   0.309              0.759
genderm     -0.29698    0.41769  -0.711              0.481
prev_expyes -0.05942    0.37589  -0.158              0.875
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.267 on 48 degrees of freedom
Multiple R-squared:  0.01647,Adjusted R-squared:  -0.045
F-statistic: 0.268 on 3 and 48 DF,  p-value: 0.8482


[[5]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-4.2399 -1.1127 -0.1532  0.8873  1.8873

Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)  5.19364    0.47088  11.030 0.00000000000000929 ***
conAware     0.21387    0.40703   0.525              0.602
genderm      0.04624    0.47015   0.098              0.922
prev_expyes -0.12717    0.42310  -0.301              0.765
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.426 on 48 degrees of freedom
Multiple R-squared:  0.006645,Adjusted R-squared:  -0.05544
F-statistic: 0.107 on 3 and 48 DF,  p-value: 0.9556
```

```
[[6]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-2.7882 -0.6078  0.2118  0.6957  1.7519

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)   5.5663     0.3993  13.941 <0.0000000000000002 ***
conAware      0.1657     0.3451   0.480               0.633
genderm       0.2218     0.3987   0.556               0.580
prev_expyes  -0.4839     0.3588  -1.349               0.184
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.209 on 48 degrees of freedom
Multiple R-squared:  0.03794,Adjusted R-squared:  -0.02219
F-statistic: 0.6309 on 3 and 48 DF,  p-value: 0.5986


[[7]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-2.8472 -0.5766  0.1528  1.1528  1.8822

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)   5.8106     0.4205  13.818 <0.0000000000000002 ***
conAware      0.5684     0.3635   1.564              0.1245
genderm      -0.2706     0.4198  -0.644              0.5223
prev_expyes  -0.6928     0.3778  -1.834              0.0729 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.273 on 48 degrees of freedom
Multiple R-squared:  0.1255,Adjusted R-squared:  0.07083
F-statistic: 2.296 on 3 and 48 DF,  p-value: 0.08959


[[8]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-4.1810 -0.3518 -0.0690  0.9667  1.9667

Coefficients:
Estimate Std. Error t value             Pr(>|t|)
(Intercept)   5.5735     0.4701  11.855 0.000000000000000724 ***
```

```
conAware      0.1477     0.4064   0.363                0.718
genderm      -0.4687     0.4694  -0.998                0.323
prev_expyes  -0.0715     0.4224  -0.169                0.866
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.424 on 48 degrees of freedom
Multiple R-squared:  0.02957,Adjusted R-squared:  -0.03108
F-statistic: 0.4875 on 3 and 48 DF,  p-value: 0.6926


[[9]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-1.3953 -1.0667 -0.2003  0.6694  3.7997

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.3474     0.3960   3.402  0.00136 **
conAware     -0.1949     0.3423  -0.569  0.57172
genderm       0.8697     0.3954   2.199  0.03270 *
prev_expyes   0.1781     0.3559   0.501  0.61898
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.199 on 48 degrees of freedom
Multiple R-squared:  0.126,Adjusted R-squared:  0.07139
F-statistic: 2.307 on 3 and 48 DF,  p-value: 0.08843


[[10]]

Call:
lm(formula = substitute(i ~ conAware + gender + prev_exp, list(i = as.name(x))),
data = d1)

Residuals:
Min      1Q  Median      3Q     Max
-2.3394 -0.3043 -0.1758  0.6694  1.7891

Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)   6.0477     0.3192  18.947 <0.0000000000000002 ***
conAware      0.1285     0.2759   0.466              0.6435
genderm      -0.8719     0.3187  -2.736              0.0087 **
prev_expyes   0.0351     0.2868   0.122              0.9031
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9665 on 48 degrees of freedom
Multiple R-squared:  0.154,Adjusted R-squared:  0.1011
F-statistic: 2.913 on 3 and 48 DF,  p-value: 0.0438
```

A.3 Chapter 5

A.3.1 Subjective Measures

No features as referent:

```
[1] "AwareIntelligent"    "AwareAuthorative"    "AwareCharismatic"    "AwareStrange"         "AwareMotivating"     "AwareL
[8] "AwareJudgmental"     "AwareLikable"        "AwareFeelMotivated"  "AwareRobotEncourage" "AwareExercFun"       "AwareF
[15] "AwareAccurate"

> subj.mod <- lapply(varList, function(x) {
+     lm(substitute(i~condAll+RobotXP+Gender, list(i = as.name(x))), data = allDF2)})
> lapply(subj.mod, summary)
[[1]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.5376 -0.7888  0.1054  0.9477  2.3054

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.1555     0.4323   9.612 1.17e-14 ***
condAllaware  -0.2729     0.3902  -0.699    0.486
condAllface   -0.1034     0.4317  -0.239    0.811
condAllfull    0.4783     0.3723   1.285    0.203
RobotXP       -0.3698     0.1682  -2.199    0.031 *
GenderMale    -0.2481     0.3158  -0.786    0.435
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.24 on 74 degrees of freedom
Multiple R-squared:  0.1372,Adjusted R-squared:  0.0789
F-statistic: 2.353 on 5 and 74 DF,  p-value: 0.04872


[[2]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2770 -0.7395  0.1919  0.7230  1.7778

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.3053     0.3593   9.198 7.03e-14 ***
condAllaware   0.3295     0.3243   1.016   0.3129
condAllface   -0.6191     0.3588  -1.726   0.0886 .
condAllfull    0.1791     0.3095   0.579   0.5647
RobotXP        0.1243     0.1398   0.889   0.3769
GenderMale    -0.4559     0.2625  -1.737   0.0866 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.031 on 74 degrees of freedom
Multiple R-squared:  0.121,Adjusted R-squared:  0.06165
F-statistic: 2.038 on 5 and 74 DF,  p-value: 0.08303
```

```
[[3]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2647 -0.5862 -0.2197  0.6657  2.5403

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.5581     0.3807   6.720 3.25e-09 ***
condAllaware   0.2486     0.3435   0.724   0.4716
condAllface    0.7667     0.3801   2.017   0.0473 *
condAllfull    0.8050     0.3278   2.455   0.0164 *
RobotXP       -0.1119     0.1481  -0.756   0.4523
GenderMale     0.1255     0.2781   0.451   0.6532
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.092 on 74 degrees of freedom
Multiple R-squared:  0.09461,Adjusted R-squared:  0.03343
F-statistic: 1.546 on 5 and 74 DF,  p-value: 0.186


[[4]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-1.7734 -0.6707  0.1091  0.4652  2.6345

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.00841    0.38993   7.715 4.47e-11 ***
condAllaware -0.35148    0.35192  -0.999    0.321
condAllface  -0.35063    0.38935  -0.901    0.371
condAllfull  -0.22480    0.33583  -0.669    0.505
RobotXP      -0.05643    0.15169  -0.372    0.711
GenderMale   -0.06566    0.28487  -0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.119 on 74 degrees of freedom
Multiple R-squared:  0.0225,Adjusted R-squared:  -0.04355
F-statistic: 0.3407 on 5 and 74 DF,  p-value: 0.8866


[[5]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
```

```
Min      1Q  Median      3Q     Max
-2.3054 -0.6720  0.2892  0.5939  1.8909

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.06389    0.35981   8.515 1.37e-12 ***
condAllaware  0.35390    0.32473   1.090    0.279
condAllface   0.51040    0.35927   1.421    0.160
condAllfull   0.34565    0.30988   1.115    0.268
RobotXP       0.04520    0.13997   0.323    0.748
GenderMale    0.06072    0.26286   0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.032 on 74 degrees of freedom
Multiple R-squared:  0.03816,Adjusted R-squared:  -0.02683
F-statistic: 0.5871 on 5 and 74 DF,  p-value: 0.7097


[[6]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.93395 -0.80448 -0.09184  0.56170  2.31210

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.0504     0.3790   8.048 1.05e-11 ***
condAllaware  -0.3043     0.3421  -0.889   0.3766
condAllface    0.2938     0.3785   0.776   0.4401
condAllfull    0.1296     0.3265   0.397   0.6925
RobotXP       -0.2460     0.1475  -1.669   0.0994 .
GenderMale     0.1988     0.2769   0.718   0.4750
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.087 on 74 degrees of freedom
Multiple R-squared:  0.05693,Adjusted R-squared:  -0.006787
F-statistic: 0.8935 on 5 and 74 DF,  p-value: 0.49


[[7]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.6509 -0.7001  0.1464  0.8626  2.0278

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.90679    0.39403   7.377 1.94e-10 ***
condAllaware  0.31702    0.35562   0.891   0.3756
condAllface   0.54772    0.39345   1.392   0.1681
condAllfull   0.78099    0.33936   2.301   0.0242 *
```

```
RobotXP       0.06546    0.15328   0.427   0.6706
GenderMale   -0.11863    0.28786  -0.412   0.6814
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.131 on 74 degrees of freedom
Multiple R-squared:  0.07957,Adjusted R-squared:  0.01737
F-statistic: 1.279 on 5 and 74 DF,  p-value: 0.2819


[[8]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.79620 -0.98119 -0.00474  0.83890  2.51992

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.818256   0.403630   6.982 1.06e-09 ***
condAllaware  0.001184   0.364279   0.003    0.997
condAllface  -0.334055   0.403029  -0.829    0.410
condAllfull   0.072129   0.347627   0.207    0.836
RobotXP      -0.094186   0.157017  -0.600    0.550
GenderMale   -0.243995   0.294871  -0.827    0.411
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.158 on 74 degrees of freedom
Multiple R-squared:  0.04014,Adjusted R-squared:  -0.02472
F-statistic: 0.6188 on 5 and 74 DF,  p-value: 0.6858


[[9]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7663 -0.6581  0.1768  0.3419  1.8298

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.14751    0.31115  10.116 1.34e-15 ***
condAllaware  0.66798    0.28082   2.379   0.0200 *
condAllface   0.51431    0.31069   1.655   0.1021
condAllfull   0.64175    0.26798   2.395   0.0192 *
RobotXP       0.02273    0.12104   0.188   0.8515
GenderMale   -0.07193    0.22731  -0.316   0.7526
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8927 on 74 degrees of freedom
Multiple R-squared:  0.0955,Adjusted R-squared:  0.03439
F-statistic: 1.563 on 5 and 74 DF,  p-value: 0.1812
```

```
[[10]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2146 -0.6253  0.1293  0.8232  1.7853

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.512882   0.372818   6.740 2.99e-09 ***
condAllaware 0.079801   0.336471   0.237    0.813
condAllface  0.609707   0.372263   1.638    0.106
condAllfull  0.298392   0.321090   0.929    0.356
RobotXP      0.029697   0.145031   0.205    0.838
GenderMale   0.002972   0.272361   0.011    0.991
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.07 on 74 degrees of freedom
Multiple R-squared:  0.04656,Adjusted R-squared:  -0.01786
F-statistic: 0.7228 on 5 and 74 DF,  p-value: 0.6085


[[11]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.6129 -0.6406  0.1902  0.6243  1.7892

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.9496     0.3999   9.876 3.75e-15 ***
condAllaware  -0.1191     0.3609  -0.330    0.742
condAllface    0.1926     0.3993   0.482    0.631
condAllfull    0.2624     0.3444   0.762    0.449
RobotXP       -0.2011     0.1556  -1.293    0.200
GenderMale    -0.1356     0.2922  -0.464    0.644
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.147 on 74 degrees of freedom
Multiple R-squared:  0.04443,Adjusted R-squared:  -0.02013
F-statistic: 0.6882 on 5 and 74 DF,  p-value: 0.6339


[[12]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
```

```
-2.4904 -0.6601  0.1336  0.7328  1.5336

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.89996    0.35023  11.136   <2e-16 ***
condAllaware  0.39996    0.31608   1.265   0.2097
condAllface   0.84643    0.34971   2.420   0.0180 *
condAllfull   0.55546    0.30163   1.842   0.0696 .
RobotXP      -0.08533    0.13624  -0.626   0.5331
GenderMale   -0.17758    0.25586  -0.694   0.4898
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.005 on 74 degrees of freedom
Multiple R-squared:  0.08828,Adjusted R-squared:  0.02668
F-statistic: 1.433 on 5 and 74 DF,  p-value: 0.2224


[[13]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8708 -1.1838 -0.1669  1.1292  2.7143

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.40533    0.47462   5.068 2.87e-06 ***
condAllaware -0.39939    0.42835  -0.932    0.354
condAllface   0.04984    0.47392   0.105    0.917
condAllfull   0.18570    0.40877   0.454    0.651
RobotXP      -0.11907    0.18463  -0.645    0.521
GenderMale    0.51792    0.34673   1.494    0.140
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.362 on 74 degrees of freedom
Multiple R-squared:  0.04702,Adjusted R-squared:  -0.01737
F-statistic: 0.7303 on 5 and 74 DF,  p-value: 0.603


[[14]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.1160 -1.4061 -0.4276  1.2607  2.5292

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.17054    0.54128   4.010 0.000144 ***
condAllaware  0.06381    0.48851   0.131 0.896432
condAllface   0.32271    0.54047   0.597 0.552268
condAllfull   0.45061    0.46617   0.967 0.336884
RobotXP       0.15013    0.21056   0.713 0.478091
```

```
GenderMale    0.17239    0.39543   0.436 0.664136
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.553 on 74 degrees of freedom
Multiple R-squared:  0.03108,Adjusted R-squared:  -0.03439
F-statistic: 0.4747 on 5 and 74 DF,  p-value: 0.794


[[15]]

Call:
lm(formula = substitute(i ~ condAll + RobotXP + Gender, list(i = as.name(x))),
data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7072 -0.5849  0.3093  0.5006  1.3033

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.9768     0.3525  11.282   <2e-16 ***
condAllaware  -0.1778     0.3181  -0.559   0.5779
condAllface    0.7997     0.3520   2.272   0.0260 *
condAllfull    0.7075     0.3036   2.330   0.0225 *
RobotXP       -0.1024     0.1371  -0.746   0.4578
GenderMale     0.1129     0.2575   0.438   0.6623
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.011 on 74 degrees of freedom
Multiple R-squared:  0.1485,Adjusted R-squared:  0.09094
F-statistic: 2.581 on 5 and 74 DF,  p-value: 0.03305
```

Aware as referent:

```
> subj.mod <- lapply(varList, function(x) {
+     lm(substitute(i~condNone+condFace+condInc+RobotXP+Gender, list(i = as.name(x))), data = allDF2)})
> lapply(subj.mod, summary)
[[1]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.5376 -0.7888  0.1054  0.9477  2.3054

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.8826     0.4458   8.709 5.89e-13 ***
condNone      0.2729     0.3902   0.699   0.4865
condFace      0.1695     0.4408   0.385   0.7017
condInc       0.7512     0.3830   1.961   0.0536 .
RobotXP      -0.3698     0.1682  -2.199   0.0310 *
GenderMale   -0.2481     0.3158  -0.786   0.4347
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.24 on 74 degrees of freedom
Multiple R-squared:  0.1372,Adjusted R-squared:  0.0789
```

```
F-statistic: 2.353 on 5 and 74 DF,  p-value: 0.04872


[[2]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2770 -0.7395  0.1919  0.7230  1.7778

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.6348     0.3706   9.809 5.01e-15 ***
condNone     -0.3295     0.3243  -1.016   0.3129
condFace     -0.9486     0.3664  -2.589   0.0116 *
condInc      -0.1504     0.3184  -0.473   0.6379
RobotXP       0.1243     0.1398   0.889   0.3769
GenderMale   -0.4559     0.2625  -1.737   0.0866 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.031 on 74 degrees of freedom
Multiple R-squared:  0.121,Adjusted R-squared:  0.06165
F-statistic: 2.038 on 5 and 74 DF,  p-value: 0.08303


[[3]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2647 -0.5862 -0.2197  0.6657  2.5403

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.8066     0.3925   7.150 5.16e-10 ***
condNone     -0.2486     0.3435  -0.724    0.472
condFace      0.5182     0.3882   1.335    0.186
condInc       0.5564     0.3372   1.650    0.103
RobotXP      -0.1119     0.1481  -0.756    0.452
GenderMale    0.1255     0.2781   0.451    0.653
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.092 on 74 degrees of freedom
Multiple R-squared:  0.09461,Adjusted R-squared:  0.03343
F-statistic: 1.546 on 5 and 74 DF,  p-value: 0.186


[[4]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)
```

```
Residuals:
Min      1Q  Median      3Q     Max
-1.7734 -0.6707  0.1091  0.4652  2.6345

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.6569301  0.4020970   6.608 5.25e-09 ***
condNone     0.3514782  0.3519190   0.999    0.321
condFace     0.0008514  0.3976135   0.002    0.998
condInc      0.1266766  0.3454546   0.367    0.715
RobotXP     -0.0564325  0.1516896  -0.372    0.711
GenderMale  -0.0656625  0.2848660  -0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.119 on 74 degrees of freedom
Multiple R-squared:  0.0225,Adjusted R-squared:  -0.04355
F-statistic: 0.3407 on 5 and 74 DF,  p-value: 0.8866


[[5]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.3054 -0.6720  0.2892  0.5939  1.8909

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.417796   0.371030   9.212 6.63e-14 ***
condNone    -0.353903   0.324729  -1.090    0.279
condFace     0.156493   0.366893   0.427    0.671
condInc     -0.008254   0.318764  -0.026    0.979
RobotXP      0.045203   0.139970   0.323    0.748
GenderMale   0.060723   0.262856   0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.032 on 74 degrees of freedom
Multiple R-squared:  0.03816,Adjusted R-squared:  -0.02683
F-statistic: 0.5871 on 5 and 74 DF,  p-value: 0.7097


[[6]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.93395 -0.80448 -0.09184  0.56170  2.31210

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.7461     0.3909   7.026  8.8e-10 ***
condNone      0.3043     0.3421   0.889   0.3766
condFace      0.5981     0.3865   1.547   0.1261
```

```
condInc       0.4339     0.3358   1.292   0.2004
RobotXP      -0.2460     0.1475  -1.669   0.0994 .
GenderMale    0.1988     0.2769   0.718   0.4750
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.087 on 74 degrees of freedom
Multiple R-squared:  0.05693,Adjusted R-squared:  -0.006787
F-statistic: 0.8935 on 5 and 74 DF,  p-value: 0.49


[[7]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.6509 -0.7001  0.1464  0.8626  2.0278

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.22381    0.40632   7.934 1.73e-11 ***
condNone    -0.31702    0.35562  -0.891    0.376
condFace     0.23070    0.40179   0.574    0.568
condInc      0.46397    0.34908   1.329    0.188
RobotXP      0.06546    0.15328   0.427    0.671
GenderMale  -0.11863    0.28786  -0.412    0.681
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.131 on 74 degrees of freedom
Multiple R-squared:  0.07957,Adjusted R-squared:  0.01737
F-statistic: 1.279 on 5 and 74 DF,  p-value: 0.2819


[[8]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.79620 -0.98119 -0.00474  0.83890  2.51992

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.819441   0.416219   6.774 2.59e-09 ***
condNone    -0.001184   0.364279  -0.003    0.997
condFace    -0.335239   0.411578  -0.815    0.418
condInc      0.070944   0.357587   0.198    0.843
RobotXP     -0.094186   0.157017  -0.600    0.550
GenderMale  -0.243995   0.294871  -0.827    0.411
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.158 on 74 degrees of freedom
Multiple R-squared:  0.04014,Adjusted R-squared:  -0.02472
F-statistic: 0.6188 on 5 and 74 DF,  p-value: 0.6858
```

```
[[9]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7663 -0.6581  0.1768  0.3419  1.8298

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.81549    0.32086  11.892   <2e-16 ***
condNone    -0.66798    0.28082  -2.379    0.020 *
condFace    -0.15366    0.31728  -0.484    0.630
condInc     -0.02623    0.27566  -0.095    0.924
RobotXP      0.02273    0.12104   0.188    0.852
GenderMale  -0.07193    0.22731  -0.316    0.753
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8927 on 74 degrees of freedom
Multiple R-squared:  0.0955,Adjusted R-squared:  0.03439
F-statistic: 1.563 on 5 and 74 DF,  p-value: 0.1812


[[10]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2146 -0.6253  0.1293  0.8232  1.7853

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.592684   0.384446   6.744 2.94e-09 ***
condNone    -0.079801   0.336471  -0.237    0.813
condFace     0.529905   0.380160   1.394    0.168
condInc      0.218591   0.330290   0.662    0.510
RobotXP      0.029697   0.145031   0.205    0.838
GenderMale   0.002972   0.272361   0.011    0.991
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.07 on 74 degrees of freedom
Multiple R-squared:  0.04656,Adjusted R-squared:  -0.01786
F-statistic: 0.7228 on 5 and 74 DF,  p-value: 0.6085


[[11]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
```

```
Min      1Q  Median      3Q     Max
-2.6129 -0.6406  0.1902  0.6243  1.7892

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.8305     0.4124   9.288 4.75e-14 ***
condNone      0.1191     0.3609   0.330    0.742
condFace      0.3117     0.4078   0.764    0.447
condInc       0.3815     0.3543   1.077    0.285
RobotXP      -0.2011     0.1556  -1.293    0.200
GenderMale   -0.1356     0.2922  -0.464    0.644
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.147 on 74 degrees of freedom
Multiple R-squared:  0.04443,Adjusted R-squared:  -0.02013
F-statistic: 0.6882 on 5 and 74 DF,  p-value: 0.6339


[[12]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.4904 -0.6601  0.1336  0.7328  1.5336

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.29992    0.36115  11.906   <2e-16 ***
condNone    -0.39996    0.31608  -1.265    0.210
condFace     0.44647    0.35712   1.250    0.215
condInc      0.15550    0.31028   0.501    0.618
RobotXP     -0.08533    0.13624  -0.626    0.533
GenderMale  -0.17758    0.25586  -0.694    0.490
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.005 on 74 degrees of freedom
Multiple R-squared:  0.08828,Adjusted R-squared:  0.02668
F-statistic: 1.433 on 5 and 74 DF,  p-value: 0.2224


[[13]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8708 -1.1838 -0.1669  1.1292  2.7143

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.0059     0.4894   4.099 0.000105 ***
condNone      0.3994     0.4284   0.932 0.354171
condFace      0.4492     0.4840   0.928 0.356315
condInc       0.5851     0.4205   1.391 0.168253
```

```
RobotXP      -0.1191     0.1846  -0.645 0.520990
GenderMale    0.5179     0.3467   1.494 0.139500
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.362 on 74 degrees of freedom
Multiple R-squared:  0.04702,Adjusted R-squared:  -0.01737
F-statistic: 0.7303 on 5 and 74 DF,  p-value: 0.603


[[14]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.1160 -1.4061 -0.4276  1.2607  2.5292

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.23435    0.55816   4.003 0.000147 ***
condNone    -0.06381    0.48851  -0.131 0.896432
condFace     0.25890    0.55193   0.469 0.640389
condInc      0.38680    0.47953   0.807 0.422465
RobotXP      0.15013    0.21056   0.713 0.478091
GenderMale   0.17239    0.39543   0.436 0.664136
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.553 on 74 degrees of freedom
Multiple R-squared:  0.03108,Adjusted R-squared:  -0.03439
F-statistic: 0.4747 on 5 and 74 DF,  p-value: 0.794


[[15]]

Call:
lm(formula = substitute(i ~ condNone + condFace + condInc + RobotXP +
Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7072 -0.5849  0.3093  0.5006  1.3033

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.7990     0.3635  10.451  3.2e-16 ***
condNone      0.1778     0.3181   0.559  0.57792
condFace      0.9775     0.3594   2.719  0.00814 **
condInc       0.8853     0.3123   2.835  0.00591 **
RobotXP      -0.1024     0.1371  -0.746  0.45778
GenderMale    0.1129     0.2575   0.438  0.66233
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.011 on 74 degrees of freedom
Multiple R-squared:  0.1485,Adjusted R-squared:  0.09094
F-statistic: 2.581 on 5 and 74 DF,  p-value: 0.03305
```

Face tracking as referent:

```
> subj.mod <- lapply(varList, function(x) {
+     lm(substitute(i~condNone+condAware+condInc+RobotXP+Gender, list(i = as.name(x))), data = allDF2)})
> lapply(subj.mod, summary)
[[1]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.5376 -0.7888  0.1054  0.9477  2.3054

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.0521      0.5376   7.538 9.65e-11 ***
condNone      0.1034      0.4317   0.239    0.811
condAware    -0.1695      0.4408  -0.385    0.702
condInc       0.5817      0.4151   1.401    0.165
RobotXP      -0.3698      0.1682  -2.199    0.031 *
GenderMale   -0.2481      0.3158  -0.786    0.435
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.24 on 74 degrees of freedom
Multiple R-squared:  0.1372,Adjusted R-squared:  0.0789
F-statistic: 2.353 on 5 and 74 DF,  p-value: 0.04872


[[2]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.2770 -0.7395  0.1919  0.7230  1.7778

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.6862      0.4468   6.012 6.41e-08 ***
condNone      0.6191      0.3588   1.726   0.0886 .
condAware     0.9486      0.3664   2.589   0.0116 *
condInc       0.7982      0.3450   2.314   0.0235 *
RobotXP       0.1243      0.1398   0.889   0.3769
GenderMale   -0.4559      0.2625  -1.737   0.0866 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.031 on 74 degrees of freedom
Multiple R-squared:  0.121,Adjusted R-squared:  0.06165
F-statistic: 2.038 on 5 and 74 DF,  p-value: 0.08303


[[3]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)
```

```
Residuals:
Min      1Q  Median      3Q     Max
-2.2647 -0.5862 -0.2197  0.6657  2.5403

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.32480    0.47334   7.024 8.86e-10 ***
condNone    -0.76673    0.38009  -2.017   0.0473 *
condAware   -0.51816    0.38816  -1.335   0.1860
condInc      0.03824    0.36548   0.105   0.9170
RobotXP     -0.11190    0.14808  -0.756   0.4523
GenderMale   0.12546    0.27809   0.451   0.6532
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.092 on 74 degrees of freedom
Multiple R-squared:  0.09461,Adjusted R-squared:  0.03343
F-statistic: 1.546 on 5 and 74 DF,  p-value: 0.186


[[4]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-1.7734 -0.6707  0.1091  0.4652  2.6345

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.6577814  0.4848766   5.481 5.58e-07 ***
condNone     0.3506268  0.3893546   0.901    0.371
condAware   -0.0008514  0.3976135  -0.002    0.998
condInc      0.1258253  0.3743808   0.336    0.738
RobotXP     -0.0564325  0.1516896  -0.372    0.711
GenderMale  -0.0656625  0.2848660  -0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.119 on 74 degrees of freedom
Multiple R-squared:  0.0225,Adjusted R-squared:  -0.04355
F-statistic: 0.3407 on 5 and 74 DF,  p-value: 0.8866


[[5]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.3054 -0.6720  0.2892  0.5939  1.8909

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.57429    0.44741   7.989 1.36e-11 ***
condNone    -0.51040    0.35927  -1.421    0.160
condAware   -0.15649    0.36689  -0.427    0.671
```

```
condInc     -0.16475    0.34545  -0.477    0.635
RobotXP      0.04520    0.13997   0.323    0.748
GenderMale   0.06072    0.26286   0.231    0.818
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.032 on 74 degrees of freedom
Multiple R-squared:  0.03816,Adjusted R-squared:  -0.02683
F-statistic: 0.5871 on 5 and 74 DF,  p-value: 0.7097


[[6]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.93395 -0.80448 -0.09184  0.56170  2.31210

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.3442     0.4713   7.095 6.53e-10 ***
condNone     -0.2938     0.3785  -0.776   0.4401
condAware    -0.5981     0.3865  -1.547   0.1261
condInc      -0.1642     0.3639  -0.451   0.6532
RobotXP      -0.2460     0.1475  -1.669   0.0994 .
GenderMale    0.1988     0.2769   0.718   0.4750
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.087 on 74 degrees of freedom
Multiple R-squared:  0.05693,Adjusted R-squared:  -0.006787
F-statistic: 0.8935 on 5 and 74 DF,  p-value: 0.49


[[7]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.6509 -0.7001  0.1464  0.8626  2.0278

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.45451    0.48997   7.050 7.92e-10 ***
condNone    -0.54772    0.39345  -1.392    0.168
condAware   -0.23070    0.40179  -0.574    0.568
condInc      0.23327    0.37831   0.617    0.539
RobotXP      0.06546    0.15328   0.427    0.671
GenderMale  -0.11863    0.28786  -0.412    0.681
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.131 on 74 degrees of freedom
Multiple R-squared:  0.07957,Adjusted R-squared:  0.01737
F-statistic: 1.279 on 5 and 74 DF,  p-value: 0.2819
```

```
[[8]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q   Median       3Q      Max
-1.79620 -0.98119 -0.00474  0.83890  2.51992

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.48420    0.50191   4.950 4.54e-06 ***
condNone     0.33405    0.40303   0.829    0.410
condAware    0.33524    0.41158   0.815    0.418
condInc      0.40618    0.38753   1.048    0.298
RobotXP     -0.09419    0.15702  -0.600    0.550
GenderMale  -0.24399    0.29487  -0.827    0.411
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.158 on 74 degrees of freedom
Multiple R-squared:  0.04014,Adjusted R-squared:  -0.02472
F-statistic: 0.6188 on 5 and 74 DF,  p-value: 0.6858


[[9]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7663 -0.6581  0.1768  0.3419  1.8298

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.66182    0.38691   9.464 2.22e-14 ***
condNone    -0.51431    0.31069  -1.655    0.102
condAware    0.15366    0.31728   0.484    0.630
condInc      0.12744    0.29874   0.427    0.671
RobotXP      0.02273    0.12104   0.188    0.852
GenderMale  -0.07193    0.22731  -0.316    0.753
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8927 on 74 degrees of freedom
Multiple R-squared:  0.0955,Adjusted R-squared:  0.03439
F-statistic: 1.563 on 5 and 74 DF,  p-value: 0.1812


[[10]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
```

```
Min       1Q  Median      3Q     Max
-2.2146 -0.6253  0.1293  0.8232  1.7853

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.122589   0.463592   6.736 3.05e-09 ***
condNone    -0.609707   0.372263  -1.638    0.106
condAware   -0.529905   0.380160  -1.394    0.168
condInc     -0.311315   0.357947  -0.870    0.387
RobotXP      0.029697   0.145031   0.205    0.838
GenderMale   0.002972   0.272361   0.011    0.991
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.07 on 74 degrees of freedom
Multiple R-squared:  0.04656,Adjusted R-squared:  -0.01786
F-statistic: 0.7228 on 5 and 74 DF,  p-value: 0.6085


[[11]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q  Median      3Q     Max
-2.6129 -0.6406  0.1902  0.6243  1.7892

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.14218    0.49729   8.330 3.08e-12 ***
condNone    -0.19258    0.39932  -0.482    0.631
condAware   -0.31171    0.40779  -0.764    0.447
condInc      0.06982    0.38396   0.182    0.856
RobotXP     -0.20109    0.15557  -1.293    0.200
GenderMale  -0.13557    0.29216  -0.464    0.644
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.147 on 74 degrees of freedom
Multiple R-squared:  0.04443,Adjusted R-squared:  -0.02013
F-statistic: 0.6882 on 5 and 74 DF,  p-value: 0.6339


[[12]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min       1Q  Median      3Q     Max
-2.4904 -0.6601  0.1336  0.7328  1.5336

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.74639    0.43550  10.899   <2e-16 ***
condNone    -0.84643    0.34971  -2.420    0.018 *
condAware   -0.44647    0.35712  -1.250    0.215
condInc     -0.29097    0.33626  -0.865    0.390
```

```
RobotXP     -0.08533    0.13624  -0.626    0.533
GenderMale  -0.17758    0.25586  -0.694    0.490
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.005 on 74 degrees of freedom
Multiple R-squared:  0.08828,Adjusted R-squared:  0.02668
F-statistic: 1.433 on 5 and 74 DF,  p-value: 0.2224


[[13]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-1.8708 -1.1838 -0.1669  1.1292  2.7143

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.45516    0.59018   4.160 8.48e-05 ***
condNone    -0.04984    0.47392  -0.105    0.917
condAware   -0.44923    0.48397  -0.928    0.356
condInc      0.13586    0.45569   0.298    0.766
RobotXP     -0.11907    0.18463  -0.645    0.521
GenderMale   0.51792    0.34673   1.494    0.140
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.362 on 74 degrees of freedom
Multiple R-squared:  0.04702,Adjusted R-squared:  -0.01737
F-statistic: 0.7303 on 5 and 74 DF,  p-value: 0.603


[[14]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.1160 -1.4061 -0.4276  1.2607  2.5292

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.4933     0.6731   3.704 0.000406 ***
condNone     -0.3227     0.5405  -0.597 0.552268
condAware    -0.2589     0.5519  -0.469 0.640389
condInc       0.1279     0.5197   0.246 0.806277
RobotXP       0.1501     0.2106   0.713 0.478091
GenderMale    0.1724     0.3954   0.436 0.664136
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.553 on 74 degrees of freedom
Multiple R-squared:  0.03108,Adjusted R-squared:  -0.03439
F-statistic: 0.4747 on 5 and 74 DF,  p-value: 0.794
```

```
[[15]]

Call:
lm(formula = substitute(i ~ condNone + condAware + condInc +
RobotXP + Gender, list(i = as.name(x))), data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7072 -0.5849  0.3093  0.5006  1.3033

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.77650    0.43833  10.897  < 2e-16 ***
condNone    -0.79967    0.35198  -2.272  0.02600 *
condAware   -0.97747    0.35944  -2.719  0.00814 **
condInc     -0.09216    0.33844  -0.272  0.78614
RobotXP     -0.10236    0.13713  -0.746  0.45778
GenderMale   0.11291    0.25752   0.438  0.66233
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.011 on 74 degrees of freedom
Multiple R-squared:  0.1485,Adjusted R-squared:  0.09094
F-statistic: 2.581 on 5 and 74 DF,  p-value: 0.03305
```

Analysis of water consumption:

```
> waterML <- lm(Water~condAll+RobotXP+Gender, data=allDF2)
> summary(waterML)

Call:
lm(formula = Water ~ condAll + RobotXP + Gender, data = allDF2)

Residuals:
Min      1Q  Median      3Q     Max
-155.24  -72.05  -13.52   66.48  294.76

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)     12.82      40.76   0.315   0.7542
condAllaware    74.60      33.60   2.220   0.0300 *
condAllface     35.37      37.02   0.955   0.3430
condAllfull     83.19      34.42   2.417   0.0186 *
RobotXP         10.35      15.41   0.672   0.5041
GenderMale      38.52      28.65   1.345   0.1835
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 100.4 on 63 degrees of freedom
(11 observations deleted due to missingness)
Multiple R-squared:  0.1372,Adjusted R-squared:  0.06867
F-statistic: 2.003 on 5 and 63 DF,  p-value: 0.09037

> describeBy(allDF2$Water,allDF2$condAll)
$`1none`
vars  n  mean    sd median trimmed mad min max range skew kurtosis    se
X1    1 17 61.76 81.41      0   55.33   0   0 220   220  0.8    -0.99 19.75

$aware
vars  n   mean    sd median trimmed   mad min max range  skew kurtosis    se
X1    1 19 139.47 92.76    150  138.82 74.13   0 290   290 -0.16    -1.13 21.28
```

```
$face
vars  n   mean    sd median trimmed    mad min max range skew kurtosis    se
X1    1 15 102.67 99.75     90   98.46 133.43   0 260   260 0.13    -1.82 25.75

$full
vars  n   mean     sd median trimmed    mad min max range skew kurtosis   se
X1    1 18 140.83 124.29    100  130.31 140.85   0 450   450 0.81    -0.19 29.3
```

Analysis of first prompt:

```
> modDrink1 <- glm(glas1~condAll+RobotXP+Gender, data=allDF2, family="binomial")
> summary(modDrink1)

Call:
glm(formula = glas1 ~ condAll + RobotXP + Gender, family = "binomial",
data = allDF2)

Deviance Residuals:
Min       1Q   Median       3Q      Max
-1.5905  -0.9580   0.7827   0.9913   1.5407

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.70806    0.78964  -0.897   0.3699
condAllaware  0.99577    0.68462   1.454   0.1458
condAllface  -0.20750    0.74651  -0.278   0.7810
condAllfull   1.45540    0.67932   2.142   0.0322 *
RobotXP       0.09288    0.30891   0.301   0.7637
GenderMale   -0.01849    0.57964  -0.032   0.9745
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 101.076  on 72  degrees of freedom
Residual deviance:  93.194  on 67  degrees of freedom
(7 observations deleted due to missingness)
AIC: 105.19

Number of Fisher Scoring iterations: 4

> logit2prob(coef(modDrink1))
(Intercept) condAllaware  condAllface  condAllfull      RobotXP   GenderMale
0.3300285    0.7302253    0.4483093    0.8108274    0.5232028    0.4953764
```

Analysis of second prompt:

```
> modDrink2 <- glm(glas2~condNone+condAware+condInc+RobotXP+Gender, data=allDF2, family="binomial")
> summary(modDrink2)

Call:
glm(formula = glas2 ~ condNone + condAware + condInc + RobotXP +
Gender, family = "binomial", data = allDF2)

Deviance Residuals:
Min       1Q   Median       3Q      Max
-1.9611  -0.8232  -0.4176   0.8475   1.8698

Coefficients:
Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept)   0.2301     1.0190   0.226   0.8213
condNone     -1.1887     0.8722  -1.363   0.1729
condAware     1.8340     0.8168   2.245   0.0247 *
condInc       1.2073     0.7410   1.629   0.1032
RobotXP      -0.2991     0.3433  -0.871   0.3836
GenderMale   -0.2407     0.6355  -0.379   0.7049
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 101.186  on 72  degrees of freedom
Residual deviance:  80.305  on 67  degrees of freedom
(7 observations deleted due to missingness)
AIC: 92.305

Number of Fisher Scoring iterations: 4

> logit2prob(coef(modDrink2))
(Intercept)     condNone    condAware      condInc      RobotXP   GenderMale
0.5572784    0.2334862    0.8622430    0.7698201    0.4257692    0.4401137
```

A.4 Chapter 6

A.4.1 Subjective Ratings

```
> varList
[1] "Likeable"       "Engaging"       "Credible"       "Enthusiastic"    "Boring"         "IntoAccount"     "RespondActions
[8] "PerceiveYou"     "warmth"         "competence"     "discomfort"
> subj.mod <- lapply(varList, function(x) {
+     lm(substitute(i~IncCond+Gender+RobotXP, list(i = as.name(x))), data = newDF1)})
> lapply(subj.mod, summary)
[[1]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-3.0457 -0.6114  0.2935  0.9023  1.5112

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)     3.3243     0.5148    6.458 5.44e-08 ***
IncCondnoInc    0.2178     0.3347    0.651    0.5184
RobotXP         0.3391     0.1977    1.715    0.0929 .
GenderMale     -0.5137     0.3670   -1.400    0.1681
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.183 on 47 degrees of freedom
Multiple R-squared:  0.09505,Adjusted R-squared:  0.03729
F-statistic: 1.646 on 3 and 47 DF,  p-value: 0.1916


[[2]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)
```

```
Residuals:
Min      1Q  Median      3Q     Max
-1.6672 -0.6400  0.1588  0.4457  1.5587

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.7166     0.3563  10.432 8.03e-14 ***
IncCondnoInc  -0.2260     0.2316  -0.976    0.334
RobotXP        0.1740     0.1368   1.272    0.210
GenderMale    -0.3974     0.2540  -1.564    0.124
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8191 on 47 degrees of freedom
Multiple R-squared:  0.07687,Adjusted R-squared:  0.01795
F-statistic: 1.305 on 3 and 47 DF,  p-value: 0.284


[[3]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min       1Q   Median       3Q      Max
-2.12139 -0.60199 -0.04336  0.72630  1.77707

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.2622     0.3997   8.162 1.46e-10 ***
IncCondnoInc   0.6173     0.2598   2.376   0.0216 *
RobotXP        0.1015     0.1535   0.662   0.5115
GenderMale    -0.2423     0.2849  -0.850   0.3994
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9189 on 47 degrees of freedom
Multiple R-squared:  0.1359,Adjusted R-squared:  0.08073
F-statistic: 2.464 on 3 and 47 DF,  p-value: 0.07397


[[4]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-2.1061 -1.1061  0.1191  1.2289  2.2441

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.89608    0.66601   4.348 7.31e-05 ***
IncCondnoInc -0.09472    0.43298  -0.219    0.828
RobotXP      -0.01515    0.25578  -0.059    0.953
GenderMale    0.24029    0.47483   0.506    0.615
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.531 on 47 degrees of freedom
Multiple R-squared:  0.00707,Adjusted R-squared:  -0.05631
F-statistic: 0.1116 on 3 and 47 DF,  p-value: 0.9529


[[5]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-1.2896 -0.7412 -0.2896  0.6175  2.6235

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.64953    0.45874   3.596 0.000774 ***
IncCondnoInc  0.26570    0.29823   0.891 0.377520
RobotXP      -0.08699    0.17618  -0.494 0.623777
GenderMale    0.63530    0.32706   1.942 0.058086 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.055 on 47 degrees of freedom
Multiple R-squared:  0.08249,Adjusted R-squared:  0.02392
F-statistic: 1.408 on 3 and 47 DF,  p-value: 0.2521


[[6]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-2.6153 -0.6831  0.1553  0.6081  1.4726

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.77688    0.44535   8.481 5.84e-11 ***
IncCondnoInc  0.15564    0.29078   0.535    0.595
RobotXP       0.06779    0.17064   0.397    0.693
GenderMale   -0.45279    0.31771  -1.425    0.161
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.02 on 46 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.05195,Adjusted R-squared:  -0.009879
F-statistic: 0.8402 on 3 and 46 DF,  p-value: 0.4789


[[7]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
```

```
Min      1Q  Median      3Q     Max
-1.8062 -0.8011  0.0346  0.5741  1.5741

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.76494    0.42297   8.901 1.44e-11 ***
IncCondnoInc  0.38032    0.27617   1.377    0.175
RobotXP      -0.02069    0.16207  -0.128    0.899
GenderMale   -0.27699    0.30175  -0.918    0.363
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9689 on 46 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.0628,Adjusted R-squared:  0.001675
F-statistic: 1.027 on 3 and 46 DF,  p-value: 0.3893


[[8]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min       1Q   Median       3Q      Max
-2.24155 -0.51761 -0.01776  0.75845  1.60560

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.9762     0.3983   9.983 7.04e-13 ***
IncCondnoInc   0.2238     0.2651   0.844   0.4032
RobotXP       -0.1529     0.1551  -0.985   0.3298
GenderMale    -0.4998     0.2941  -1.700   0.0962 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9122 on 44 degrees of freedom
(3 observations deleted due to missingness)
Multiple R-squared:  0.1172,Adjusted R-squared:  0.05705
F-statistic: 1.948 on 3 and 44 DF,  p-value: 0.1358


[[9]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min       1Q   Median       3Q      Max
-1.86717 -0.69960  0.04823  0.56787  1.81787

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.73793    0.41114   6.659 2.69e-08 ***
IncCondnoInc -0.05146    0.26729  -0.193    0.848
RobotXP       0.09864    0.15790   0.625    0.535
GenderMale   -0.25307    0.29313  -0.863    0.392
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.9452 on 47 degrees of freedom
Multiple R-squared:  0.01992,Adjusted R-squared:  -0.04264
F-statistic: 0.3184 on 3 and 47 DF,  p-value: 0.812


[[10]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min       1Q   Median       3Q      Max
-1.61068 -0.34082  0.05599  0.40637  1.37883

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.86626    0.31927  12.110 4.69e-16 ***
IncCondnoInc  0.17716    0.20756   0.854    0.398
RobotXP       0.01704    0.12262   0.139    0.890
GenderMale   -0.47338    0.22763  -2.080    0.043 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.734 on 47 degrees of freedom
Multiple R-squared:  0.1062,Adjusted R-squared:  0.04917
F-statistic: 1.862 on 3 and 47 DF,  p-value: 0.149


[[11]]

Call:
lm(formula = substitute(i ~ IncCond + RobotXP + Gender, list(i = as.name(x))),
data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-0.9118 -0.4394 -0.1889  0.2778  1.7549

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.78927    0.27476   6.512  4.5e-08 ***
IncCondnoInc  0.03471    0.17863   0.194    0.847
RobotXP      -0.04502    0.10552  -0.427    0.672
GenderMale    0.17782    0.19589   0.908    0.369
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6317 on 47 degrees of freedom
Multiple R-squared:  0.01854,Adjusted R-squared:  -0.04411
F-statistic: 0.2959 on 3 and 47 DF,  p-value: 0.8281
```

A.4.2 Interaction with Gender

```
> respMod <- lm(warmth~Gender*IncCond+RobotXP, data=newDF1)
> summary(respMod)

Call:
lm(formula = warmth ~ Gender * IncCond + RobotXP, data = newDF1)

Residuals:
```

```
Min       1Q   Median       3Q      Max
-1.86761 -0.56085 -0.02079  0.52882  1.91258

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)              2.26905    0.44221   5.131 5.65e-06 ***
GenderMale               0.42677    0.40571   1.052   0.2983
IncCondnoInc             0.82739    0.45724   1.810   0.0769 .
RobotXP                  0.08589    0.15113   0.568   0.5726
GenderMale:IncCondnoInc -1.27425    0.54964  -2.318   0.0249 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9041 on 46 degrees of freedom
Multiple R-squared:  0.1224,Adjusted R-squared:  0.04614
F-statistic: 1.605 on 4 and 46 DF,  p-value: 0.1892

> respMod <- lm(competence~IncCond*Gender+RobotXP, data=newDF1)
> summary(respMod)

Call:
lm(formula = competence ~ IncCond * Gender + RobotXP, data = newDF1)

Residuals:
Min       1Q   Median       3Q      Max
-1.67508 -0.35866 -0.00841  0.44822  1.46183

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)              3.69145    0.35850  10.297 1.59e-13 ***
IncCondnoInc             0.50481    0.37069   1.362    0.180
GenderMale              -0.21992    0.32891  -0.669    0.507
RobotXP                  0.01229    0.12252   0.100    0.921
IncCondnoInc:GenderMale -0.47506    0.44560  -1.066    0.292
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.733 on 46 degrees of freedom
Multiple R-squared:  0.1278,Adjusted R-squared:  0.05193
F-statistic: 1.685 on 4 and 46 DF,  p-value: 0.1698

> respMod <- lm(discomfort~IncCond*Gender+RobotXP, data=newDF1)
> summary(respMod)

Call:
lm(formula = discomfort ~ IncCond * Gender + RobotXP, data = newDF1)

Residuals:
Min      1Q  Median      3Q     Max
-0.9276 -0.4276 -0.1827  0.3024  1.8121

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)               1.6615     0.3096   5.367 2.54e-06 ***
IncCondnoInc              0.2742     0.3201   0.857    0.396
GenderMale                0.3631     0.2840   1.278    0.208
RobotXP                  -0.0485     0.1058  -0.458    0.649
IncCondnoInc:GenderMale  -0.3473     0.3848  -0.902    0.372
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Residual standard error: 0.6329 on 46 degrees of freedom
Multiple R-squared:  0.03562,Adjusted R-squared:  -0.04824
F-statistic: 0.4247 on 4 and 46 DF,  p-value: 0.79
```

A.4.3 Interaction Between Perception and Performance

```
> varList2
[1] "Likeable"      "Engaging"      "Credible"      "Enthusiastic"   "Boring"        "IntoAccount"     "RespondActions
[8] "PerceiveYou"   "warmth"        "competence"    "discomfort"

> subj.mod <- lapply(varList, function(x) {
+     lm(substitute(i~IncCond+value*Gender+RobotXP, list(i = as.name(x))), data = incMan)})
> lapply(subj.mod, summary)
[[1]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min      1Q  Median      3Q     Max
-2.9573 -0.5899  0.3284  1.0619  1.4477

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.682878   0.494404   7.449 1.32e-10 ***
IncCondnoInc      0.083510   0.272691   0.306   0.7603
value            -0.010947   0.022248  -0.492   0.6241
GenderMale       -0.641587   0.490052  -1.309   0.1945
RobotXP           0.283811   0.156332   1.815   0.0735 .
value:GenderMale  0.008313   0.026185   0.317   0.7518
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.166 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.06941,Adjusted R-squared:  0.00737
F-statistic: 1.119 on 5 and 75 DF,  p-value: 0.3578


[[2]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.51375 -0.51488 -0.02138  0.49624  1.68957

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.677118   0.350127  10.502   <2e-16 ***
IncCondnoInc     -0.195745   0.193114  -1.014    0.314
value             0.003235   0.015755   0.205    0.838
GenderMale       -0.510825   0.347045  -1.472    0.145
RobotXP           0.162209   0.110712   1.465    0.147
value:GenderMale -0.001624   0.018543  -0.088    0.930
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8255 on 75 degrees of freedom
```

```
(21 observations deleted due to missingness)
Multiple R-squared:  0.09731,Adjusted R-squared:  0.03713
F-statistic: 1.617 on 5 and 75 DF,  p-value: 0.1659


[[3]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.92411 -0.70166 -0.01476  0.73204  1.76774

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.069667   0.375179   8.182 5.38e-12 ***
IncCondnoInc      0.579907   0.206932   2.802  0.00645 **
value             0.012353   0.016883   0.732  0.46662
GenderMale       -0.363986   0.371877  -0.979  0.33084
RobotXP           0.155067   0.118633   1.307  0.19517
value:GenderMale -0.002453   0.019870  -0.123  0.90208
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8846 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.1949,Adjusted R-squared:  0.1412
F-statistic: 3.631 on 5 and 75 DF,  p-value: 0.005378


[[4]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min      1Q  Median      3Q     Max
-2.1860 -1.1156  0.1188  1.0328  2.6996

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       2.65985    0.64079   4.151 8.66e-05 ***
IncCondnoInc      0.03269    0.35343   0.092    0.927
value             0.01471    0.02883   0.510    0.612
GenderMale        0.56788    0.63515   0.894    0.374
RobotXP           0.01065    0.20262   0.053    0.958
value:GenderMale -0.03181    0.03394  -0.937    0.352
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.511 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.01618,Adjusted R-squared:  -0.04941
F-statistic: 0.2466 on 5 and 75 DF,  p-value: 0.9403


[[5]]
```

```
Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min      1Q  Median      3Q     Max
-1.4647 -0.7339 -0.2377  0.5665  2.4738

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.928644   0.441574   4.368 3.97e-05 ***
IncCondnoInc      0.339445   0.243552   1.394    0.168
value            -0.012999   0.019870  -0.654    0.515
GenderMale        0.680833   0.437687   1.556    0.124
RobotXP          -0.106557   0.139627  -0.763    0.448
value:GenderMale -0.006589   0.023387  -0.282    0.779
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.041 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.1037,Adjusted R-squared:  0.0439
F-statistic: 1.735 on 5 and 75 DF,  p-value: 0.1371


[[6]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-2.55141 -0.62203  0.06562  0.71580  1.54298

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.07172    0.43530   9.354 4.05e-14 ***
IncCondnoInc      0.12737    0.24173   0.527   0.5999
value            -0.02007    0.01956  -1.027   0.3080
GenderMale       -0.90471    0.43356  -2.087   0.0404 *
RobotXP           0.08749    0.13775   0.635   0.5273
value:GenderMale  0.02774    0.02305   1.204   0.2326
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.024 on 73 degrees of freedom
(23 observations deleted due to missingness)
Multiple R-squared:  0.0699,Adjusted R-squared:  0.006199
F-statistic: 1.097 on 5 and 73 DF,  p-value: 0.3693


[[7]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.92117 -0.82330  0.06359  0.68503  1.65760
```

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.813691   0.403623   9.449  2.7e-14 ***
IncCondnoInc      0.522741   0.224139   2.332   0.0225 *
value            -0.007427   0.018132  -0.410   0.6833
GenderMale       -0.450819   0.402010  -1.121   0.2658
RobotXP           0.020324   0.127724   0.159   0.8740
value:GenderMale  0.004401   0.021372   0.206   0.8374
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9495 on 73 degrees of freedom
(23 observations deleted due to missingness)
Multiple R-squared:  0.1094,Adjusted R-squared:  0.04839
F-statistic: 1.793 on 5 and 73 DF,  p-value: 0.1249


[[8]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.62961 -0.48178 -0.05875  0.63899  1.69828

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.32434    0.34882  12.397   <2e-16 ***
IncCondnoInc      0.38395    0.19565   1.962   0.0536 .
value            -0.03380    0.01567  -2.157   0.0343 *
GenderMale       -0.84772    0.34882  -2.430   0.0176 *
RobotXP          -0.11956    0.11073  -1.080   0.2838
value:GenderMale  0.03186    0.01848   1.724   0.0889 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8204 on 72 degrees of freedom
(24 observations deleted due to missingness)
Multiple R-squared:  0.1573,Adjusted R-squared:  0.09874
F-statistic: 2.687 on 5 and 72 DF,  p-value: 0.02776


[[9]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.81196 -0.64809  0.00312  0.59264  2.12847

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       2.36643    0.39311   6.020 5.97e-08 ***
IncCondnoInc      0.10366    0.21682   0.478    0.634
value             0.02066    0.01769   1.168    0.247
GenderMale        0.16549    0.38965   0.425    0.672
```

```
RobotXP           0.08747    0.12430   0.704    0.484
value:GenderMale -0.03261    0.02082  -1.566    0.122
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9268 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.06673,Adjusted R-squared:  0.004508
F-statistic: 1.072 on 5 and 75 DF,  p-value: 0.3825


[[10]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min       1Q   Median       3Q      Max
-1.58815 -0.39379  0.07027  0.46080  1.43257

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.8259573  0.3213747  11.905   <2e-16 ***
IncCondnoInc      0.2302335  0.1772557   1.299    0.198
value            -0.0004129  0.0144615  -0.029    0.977
GenderMale       -0.4771091  0.3185459  -1.498    0.138
RobotXP           0.0315405  0.1016199   0.310    0.757
value:GenderMale -0.0014193  0.0170206  -0.083    0.934
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7577 on 75 degrees of freedom
(21 observations deleted due to missingness)
Multiple R-squared:  0.1179,Adjusted R-squared:  0.05905
F-statistic: 2.004 on 5 and 75 DF,  p-value: 0.08771


[[11]]

Call:
lm(formula = substitute(i ~ IncCond + value * Gender + RobotXP,
list(i = as.name(x))), data = incMan)

Residuals:
Min      1Q  Median      3Q     Max
-0.9769 -0.4899 -0.1729  0.3077  1.6999

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.829153   0.273148   6.697 3.42e-09 ***
IncCondnoInc      0.064663   0.150656   0.429    0.669
value            -0.002023   0.012291  -0.165    0.870
GenderMale        0.231728   0.270744   0.856    0.395
RobotXP          -0.059266   0.086371  -0.686    0.495
value:GenderMale -0.001740   0.014466  -0.120    0.905
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.644 on 75 degrees of freedom
(21 observations deleted due to missingness)
```

```
Multiple R-squared:  0.02711,Adjusted R-squared:  -0.03775
F-statistic: 0.418 on 5 and 75 DF,  p-value: 0.8348
```

A.5 Chapter 7

A.5.1 Manipulation Check

```
> knowtarget <- table(thirdhand2$KnowTargetNum,thirdhand2$conGaze)
> knowtarget

ProActive Tracking
1         18         9
2         18        22
3         10         5
> chisq.test(knowtarget)

Pearson's Chi-squared test

data:  knowtarget
X-squared = 3.9052, df = 2, p-value = 0.1419
```

A.5.2 Questionnaire

```
> varList
[1] "ctrlPerfMeRobot"           "IncompetentCompetent"     "IgnorantKnowledgeable"    "UnpredictablePredictable" "Irrespon
[6] "UnintelligentIntelligent" "FoolishSensible"          "CollaboratioSuccess"      "respPerfMeRobot"          "GuessNex
> gaze <- lapply(varList, function(x) {
+     lm(substitute(i~conGaze+RobotXP+Sex, list(i = as.name(x))), data = thirdhand2)})
> lapply(gaze, summary)
[[1]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min       1Q   Median       3Q      Max
-2.92944 -0.88060  0.09705  1.14589  3.15003

Coefficients:
Estimate Std. Error t value         Pr(>|t|)
(Intercept)      3.814691   0.486269   7.845 0.0000000000175 ***
conGazeTracking 0.004140   0.370720   0.011            0.991
RobotXP         0.026491   0.224082   0.118            0.906
SexMale         0.008791   0.403879   0.022            0.983
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.644 on 79 degrees of freedom
Multiple R-squared:  0.0002247,Adjusted R-squared:  -0.03774
F-statistic: 0.005919 on 3 and 79 DF,  p-value: 0.9994


[[2]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min       1Q   Median       3Q      Max
```

```
-3.12321 -0.31927 -0.04203  0.84309  2.04483

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)      5.23242    0.33210  15.756 <0.0000000000000002 ***
conGazeTracking -0.19606    0.25318  -0.774               0.441
RobotXP         -0.08118    0.15304  -0.530               0.597
SexMale          0.16804    0.27583   0.609               0.544
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.123 on 79 degrees of freedom
Multiple R-squared:  0.01208,Adjusted R-squared:  -0.02543
F-statistic: 0.3221 on 3 and 79 DF,  p-value: 0.8094


[[3]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-3.9716 -1.0782  0.1849  1.1599  2.6740

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)       4.4326     0.4655   9.523 0.00000000000000931 ***
conGazeTracking   0.2632     0.3549   0.742               0.461
RobotXP          -0.1066     0.2145  -0.497               0.621
SexMale           0.4891     0.3866   1.265               0.210
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.574 on 79 degrees of freedom
Multiple R-squared:  0.02899,Adjusted R-squared:  -0.007881
F-statistic: 0.7863 on 3 and 79 DF,  p-value: 0.5051


[[4]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-3.1989 -0.6359  0.4007  0.5810  1.8856

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)      5.02991    0.35459  14.185 <0.0000000000000002 ***
conGazeTracking  0.09587    0.27033   0.355               0.724
RobotXP          0.08450    0.16340   0.517               0.607
SexMale          0.30457    0.29451   1.034               0.304
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.199 on 79 degrees of freedom
Multiple R-squared:  0.02337,Adjusted R-squared:  -0.01372
```

```
F-statistic:  0.63 on 3 and 79 DF,  p-value: 0.5978


[[5]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-3.3376 -1.1853  0.4815  0.7859  1.8434

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)      5.57591    0.38272  14.569 <0.0000000000000002 ***
conGazeTracking -0.12347    0.29178  -0.423               0.673
RobotXP         -0.05744    0.17636  -0.326               0.746
SexMale         -0.18094    0.31787  -0.569               0.571
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.294 on 79 degrees of freedom
Multiple R-squared:  0.009151,Adjusted R-squared:  -0.02848
F-statistic: 0.2432 on 3 and 79 DF,  p-value: 0.8659


[[6]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-3.8998 -0.8398  0.2913  1.2913  2.6649

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)       5.1221     0.4638  11.043 <0.0000000000000002 ***
conGazeTracking  -0.1910     0.3536  -0.540               0.591
RobotXP          -0.2224     0.2137  -1.040               0.301
SexMale          -0.1199     0.3852  -0.311               0.756
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.568 on 79 degrees of freedom
Multiple R-squared:  0.02015,Adjusted R-squared:  -0.01706
F-statistic: 0.5416 on 3 and 79 DF,  p-value: 0.6552


[[7]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-4.1310 -0.7750 -0.0432  0.9568  2.5810
```

```
Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)       4.2179     0.4232   9.966 0.00000000000000128 ***
conGazeTracking   0.4899     0.3227   1.518                0.133
RobotXP           0.2011     0.1950   1.031                0.306
SexMale           0.2221     0.3515   0.632                0.529
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.431 on 79 degrees of freedom
Multiple R-squared:  0.04625,Adjusted R-squared:  0.01004
F-statistic: 1.277 on 3 and 79 DF,  p-value: 0.288


[[8]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min       1Q   Median       3Q      Max
-2.91901 -0.86360  0.08099  1.07762  2.53348

Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)     4.464271   0.381385  11.705 <0.0000000000000002 ***
conGazeTracking 0.115316   0.290758   0.397                0.693
RobotXP         0.002247   0.175749   0.013                0.990
SexMale         0.337178   0.316766   1.064                0.290
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.289 on 79 degrees of freedom
Multiple R-squared:  0.01764,Adjusted R-squared:  -0.01966
F-statistic: 0.4729 on 3 and 79 DF,  p-value: 0.7021


[[9]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-2.7667 -1.0429 -0.1749  0.9571  2.9571

Coefficients:
Estimate Std. Error t value          Pr(>|t|)
(Intercept)       3.7765     0.4241   8.905 0.000000000000149 ***
conGazeTracking  -0.2762     0.3233  -0.854             0.396
RobotXP          -0.2959     0.1954  -1.514             0.134
SexMale           0.5623     0.3522   1.596             0.114
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.434 on 79 degrees of freedom
Multiple R-squared:  0.04857,Adjusted R-squared:  0.01244
F-statistic: 1.344 on 3 and 79 DF,  p-value: 0.266
```

```
[[10]]

Call:
lm(formula = substitute(i ~ conGaze + RobotXP + Sex, list(i = as.name(x))),
data = thirdhand2)

Residuals:
Min      1Q  Median      3Q     Max
-4.3930 -0.7039  0.5714  0.7333  2.2427

Coefficients:
Estimate Std. Error t value             Pr(>|t|)
(Intercept)      4.79291    0.47219  10.150 0.000000000000000565 ***
conGazeTracking  0.07282    0.35998   0.202                0.840
RobotXP         -0.03562    0.21759  -0.164                0.870
SexMale          0.59852    0.39218   1.526                0.131
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.596 on 79 degrees of freedom
Multiple R-squared:  0.03049,Adjusted R-squared:  -0.006329
F-statistic: 0.8281 on 3 and 79 DF,  p-value: 0.4823
```

A.5.3 Initial Gaze

```
> lookwhereStart

ProActive Tracking
face        112       83
other        68       54
> chisq.test(lookwhereStart)

Pearson's Chi-squared test with Yates' continuity correction

data:  lookwhereStart
X-squared = 0.032567, df = 1, p-value = 0.8568
```

A.5.4 Objective Measure

```
> overlap.mod <- lmer(overlapDurPerc~conGaze+LookInit+Sex+RobotXP+order+(1|PID), data=overlapMergedFull3)
> summary(overlap.mod)
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: overlapDurPerc ~ conGaze + LookInit + Sex + RobotXP + order +      (1 | PID)
Data: overlapMergedFull3

REML criterion at convergence: 2377.6

Scaled residuals:
Min      1Q  Median      3Q     Max
-3.3740 -0.5094  0.0038  0.4609  2.7147

Random effects:
Groups   Name        Variance Std.Dev.
PID      (Intercept) 797.4    28.24
Residual             463.1    21.52
Number of obs: 253, groups:  PID, 74

Fixed effects:
Estimate Std. Error      df t value Pr(>|t|)
(Intercept)       81.160     11.132 101.070   7.291 7.04e-11 ***
```

```
conGazeTracking   -8.121      7.327  69.720  -1.108 0.271521
LookInitface     -11.119      4.068 223.400  -2.733 0.006773 **
SexMale           23.703      8.250  72.080   2.873 0.005337 **
RobotXP          -17.016      4.311  70.870  -3.947 0.000184 ***
order             -3.340      1.269 181.830  -2.632 0.009218 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
(Intr) cnGzTr LkIntf SexMal RobtXP
conGzTrckng -0.381
LookInitfac -0.325 -0.010
SexMale     -0.371 -0.078  0.100
RobotXP     -0.600  0.193 -0.001 -0.247
order       -0.321  0.004  0.048  0.007  0.001
```

A.5.5 Interactions

```
> overlapW.mod <- lmer(overlapDurPerc~conGaze*LookInit+RobotXP+order+(1|PID), data=overlapWom)
> summary(overlapW.mod)
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: overlapDurPerc ~ conGaze * LookInit + RobotXP + order + (1 |      PID)
Data: overlapWom

REML criterion at convergence: 593

Scaled residuals:
Min       1Q  Median      3Q     Max
-2.0892 -0.3765 -0.0686  0.3729  2.6768

Random effects:
Groups   Name        Variance Std.Dev.
PID      (Intercept) 343.3    18.53
Residual             376.1    19.39
Number of obs: 69, groups:  PID, 21

Fixed effects:
Estimate Std. Error      df t value Pr(>|t|)
(Intercept)                 121.149     16.781  35.090   7.220 1.97e-08 ***
conGazeTracking             -66.258     19.604  61.270  -3.380  0.00127 **
LookInitface                -46.687     10.758  61.530  -4.340 5.40e-05 ***
RobotXP                     -15.549      7.343  18.880  -2.117  0.04772 *
order                        -5.294      2.210  48.650  -2.395  0.02050 *
conGazeTracking:LookInitface   46.220     19.001  58.860   2.432  0.01806 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
(Intr) cnGzTr LkIntf RobtXP order
conGzTrckng -0.424
LookInitfac -0.577  0.435
RobotXP     -0.657  0.123  0.021
order       -0.400  0.008  0.143 -0.023
cnGzTrck:LI  0.333 -0.867 -0.561 -0.046 -0.041
> overlapM.mod <- lmer(overlapDurPerc~conGaze*LookInit+RobotXP+order+(1|PID), data=overlapMen)
> summary(overlapM.mod)
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: overlapDurPerc ~ conGaze * LookInit + RobotXP + order + (1 |      PID)
Data: overlapMen
```

```
REML criterion at convergence: 1729.8

Scaled residuals:
Min      1Q  Median      3Q     Max
-3.2397 -0.4832  0.0751  0.4681  2.6464

Random effects:
Groups   Name        Variance Std.Dev.
PID      (Intercept) 882.5    29.71
Residual             490.6    22.15
Number of obs: 184, groups:  PID, 53

Fixed effects:
Estimate Std. Error      df t value Pr(>|t|)
(Intercept)                   96.354     13.464  78.300   7.156  3.9e-10 ***
conGazeTracking                2.135     11.013  89.850   0.194  0.84673
LookInitface                  -5.817      6.118 155.250  -0.951  0.34318
RobotXP                      -16.860      4.996  50.630  -3.375  0.00143 **
order                         -2.923      1.529 130.720  -1.911  0.05815 .
conGazeTracking:LookInitface  -3.656      9.241 157.920  -0.396  0.69292
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Intr) cnGzTr LkIntf RobtXP order
conGzTrckng -0.537
LookInitfac -0.362  0.385
RobotXP     -0.781  0.223  0.064
order       -0.322  0.026  0.028  0.010
cnGzTrck:LI  0.281 -0.568 -0.665 -0.095 -0.019
```

A.5.6 Discussion

```
> table(overlapMergedFull$LookatFirst,overlapMergedFull$Sex)

Female Male
face      67  116
other     28  107

> chisq.test(table(overlapMergedFull$LookatFirst,overlapMergedFull$Sex))

Pearson's Chi-squared test with Yates' continuity correction

data:  table(overlapMergedFull$LookatFirst, overlapMergedFull$Sex)
X-squared = 8.5991, df = 1, p-value = 0.003363
```

A.6 Chapter 8

A.6.1 Questionnaire

```
 [1] "ctrlPerfMeRobot"           "IncompetentCompetent"     "IgnorantKnowledgeable"     "UnpredictablePredictable"
[5] "IrresponsibleResponsible" "UnintelligentIntelligent" "CompliantNoncompliant"     "FoolishSensible"
[9] "CollaboratioSuccess"      "eagerness"                "compliance"                "respPerfMeRobot"
[13] "easyToCorrect"

> repair <- lapply(varList, function(x) {
+     lm(substitute(i~ConRepair+Sex+RobotXP, list(i = as.name(x))), data = thirdhandRepair)})
> lapply(repair, summary)
[[1]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
```

```
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.1376 -1.1560 -0.1376  1.3303  3.3761

Coefficients:
Estimate              Std. Error t value      Pr(>|t|)
(Intercept)      4.66972477064220026222  0.67615515056016350925   6.906 0.00000000638 ***
ConRepairRepair -0.53211009174311951764  0.45133591475768469747  -1.179         0.244
SexMale         -0.51376146788990750824  0.48228613176486340164  -1.065         0.292
RobotXP         -0.00000000000000006518  0.27539056176129972364   0.000         1.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.642 on 53 degrees of freedom
Multiple R-squared:  0.04108,Adjusted R-squared:  -0.0132
F-statistic: 0.7568 on 3 and 53 DF,  p-value: 0.5234


[[2]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min       1Q   Median       3Q      Max
-2.34619 -0.34619  0.04506  0.65381  2.04506

Coefficients:
Estimate Std. Error t value            Pr(>|t|)
(Intercept)       4.9982     0.3876  12.894 <0.0000000000000002 ***
ConRepairRepair   0.2032     0.2587   0.785               0.436
SexMale           0.3913     0.2765   1.415               0.163
RobotXP          -0.2464     0.1579  -1.561               0.124
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9412 on 53 degrees of freedom
Multiple R-squared:  0.07902,Adjusted R-squared:  0.02689
F-statistic: 1.516 on 3 and 53 DF,  p-value: 0.2211


[[3]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.7200 -0.8020  0.2800  0.8428  2.3330

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)       4.6236     0.6154   7.514 0.000000000673 ***
ConRepairRepair   0.4806     0.4107   1.170         0.2472
SexMale           0.4903     0.4389   1.117         0.2691
RobotXP          -0.4373     0.2506  -1.745         0.0868 .
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.494 on 53 degrees of freedom
Multiple R-squared:  0.09129,Adjusted R-squared:  0.03985
F-statistic: 1.775 on 3 and 53 DF,  p-value: 0.1632


[[4]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-2.7746 -0.7746  0.2253  0.5387  2.2254

Coefficients:
Estimate Std. Error t value        Pr(>|t|)
(Intercept)       3.9472     0.4567   8.642 0.0000000000107 ***
ConRepairRepair   0.6867     0.3049   2.252          0.0285 *
SexMale           0.6868     0.3258   2.108          0.0398 *
RobotXP           0.1407     0.1860   0.756          0.4529
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.109 on 53 degrees of freedom
Multiple R-squared:  0.1486,Adjusted R-squared:  0.1004
F-statistic: 3.083 on 3 and 53 DF,  p-value: 0.03506


[[5]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.2011 -1.1433 -0.0591  0.8000  1.8982

Coefficients:
Estimate Std. Error t value           Pr(>|t|)
(Intercept)      5.35710    0.50525  10.603 0.0000000000000104 ***
ConRepairRepair  0.04264    0.33725   0.126              0.900
SexMale         -0.05780    0.36038  -0.160              0.873
RobotXP         -0.09932    0.20578  -0.483              0.631
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.227 on 53 degrees of freedom
Multiple R-squared:  0.005945,Adjusted R-squared:  -0.05032
F-statistic: 0.1057 on 3 and 53 DF,  p-value: 0.9564


[[6]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)
```

```
Residuals:
Min      1Q  Median      3Q     Max
-4.0208 -0.8904  0.2527  0.9792  2.5624

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)       4.9160     0.6301   7.802 0.000000000232 ***
ConRepairRepair   0.3784     0.4206   0.900          0.372
SexMale          -0.2048     0.4494  -0.456          0.650
RobotXP          -0.2736     0.2566  -1.066          0.291
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.53 on 53 degrees of freedom
Multiple R-squared:  0.04682,Adjusted R-squared:  -0.007134
F-statistic: 0.8678 on 3 and 53 DF,  p-value: 0.4636


[[7]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.0699 -0.6188  0.3812  0.8154  3.3812

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)       4.2148     0.5609   7.514 0.000000000671 ***
ConRepairRepair  -0.7407     0.3744  -1.978         0.0531 .
SexMale           0.2894     0.4001   0.723         0.4726
RobotXP          -0.1447     0.2284  -0.634         0.5291
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.362 on 53 degrees of freedom
Multiple R-squared:  0.08559,Adjusted R-squared:  0.03383
F-statistic: 1.654 on 3 and 53 DF,  p-value: 0.1881


[[8]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.5546 -0.8574 -0.0525  1.0169  2.7593

Coefficients:
Estimate Std. Error t value    Pr(>|t|)
(Intercept)       3.5435     0.5792   6.118 0.000000117 ***
ConRepairRepair   0.3832     0.3866   0.991      0.3260
SexMale           0.8118     0.4131   1.965      0.0546 .
RobotXP           0.3139     0.2359   1.331      0.1890
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Residual standard error: 1.406 on 53 degrees of freedom
Multiple R-squared:  0.1128,Adjusted R-squared:  0.06257
F-statistic: 2.246 on 3 and 53 DF,  p-value: 0.09362


[[9]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min       1Q   Median       3Q      Max
-2.20784 -0.67993  0.07013  0.79216  2.63875

Coefficients:
Estimate Std. Error t value          Pr(>|t|)
(Intercept)       4.1113     0.4344   9.463 0.000000000000558 ***
ConRepairRepair   0.5279     0.2900   1.820           0.07434 .
SexMale           0.8466     0.3099   2.732           0.00853 **
RobotXP          -0.2780     0.1769  -1.571           0.12215
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.055 on 53 degrees of freedom
Multiple R-squared:  0.1813,Adjusted R-squared:  0.1349
F-statistic: 3.912 on 3 and 53 DF,  p-value: 0.01348


[[10]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.7147 -0.7147 -0.1332  0.6338  2.2853

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)       4.1368     0.4858   8.515 0.000000000017 ***
ConRepairRepair   0.7670     0.3243   2.365         0.0217 *
SexMale           0.7521     0.3465   2.170         0.0345 *
RobotXP          -0.1742     0.1979  -0.881         0.3825
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.18 on 53 degrees of freedom
Multiple R-squared:  0.1579,Adjusted R-squared:  0.1102
F-statistic: 3.312 on 3 and 53 DF,  p-value: 0.02686


[[11]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
```

```
-4.0666 -0.5733 -0.0666  0.9334  2.3165

Coefficients:
Estimate Std. Error t value       Pr(>|t|)
(Intercept)       4.1075     0.5360   7.663 0.000000000387 ***
ConRepairRepair   0.8431     0.3578   2.357         0.0222 *
SexMale           0.6502     0.3823   1.701         0.0949 .
RobotXP          -0.2671     0.2183  -1.224         0.2265
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.301 on 53 degrees of freedom
Multiple R-squared:  0.1471,Adjusted R-squared:  0.09886
F-statistic: 3.048 on 3 and 53 DF,  p-value: 0.03651


[[12]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.4081 -0.7477 -0.0053  0.9947  2.5919

Coefficients:
Estimate Std. Error t value      Pr(>|t|)
(Intercept)       4.4456     0.6137   7.244 0.00000000182 ***
ConRepairRepair  -0.6604     0.4096  -1.612         0.113
SexMale           0.3337     0.4377   0.762         0.449
RobotXP          -0.3712     0.2499  -1.485         0.143
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.49 on 53 degrees of freedom
Multiple R-squared:  0.08765,Adjusted R-squared:  0.036
F-statistic: 1.697 on 3 and 53 DF,  p-value: 0.1788


[[13]]

Call:
lm(formula = substitute(i ~ ConRepair + Sex + RobotXP, list(i = as.name(x))),
data = thirdhandRepair)

Residuals:
Min      1Q  Median      3Q     Max
-3.8741 -1.3410  0.1259  1.1594  4.0688

Coefficients:
Estimate Std. Error t value  Pr(>|t|)
(Intercept)      3.25047    0.69518   4.676 0.00002056 ***
ConRepairRepair  2.53314    0.46404   5.459 0.00000129 ***
SexMale         -0.08982    0.49586  -0.181      0.857
RobotXP         -0.40983    0.28314  -1.447      0.154
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.688 on 53 degrees of freedom
Multiple R-squared:  0.3931,Adjusted R-squared:  0.3588
```

```
F-statistic: 11.44 on 3 and 53 DF,  p-value: 0.00000678
```

A.6.2 Order Effect

```
> modHor <- lmer(value~order+(1|PID), data=splitHo)
> summary(modHor)
Linear mixed model fit by REML
t-tests use  Satterthwaite approximations to degrees of freedom ['lmerMod']
Formula: value ~ order + (1 | PID)
Data: splitHo

REML criterion at convergence: 1945.6

Scaled residuals:
Min      1Q  Median      3Q     Max
-2.4002 -0.3040 -0.1150  0.0772  8.2651

Random effects:
Groups   Name        Variance Std.Dev.
PID      (Intercept)  8.106   2.847
Residual             13.117   3.622
Number of obs: 340, groups:  PID, 85

Fixed effects:
Estimate Std. Error       df t value    Pr(>|t|)
(Intercept)    2.9240      0.5717 296.6800    5.115 0.000000565 ***
order         -0.5053      0.1757 254.0000  -2.876     0.00436 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
(Intr)
order -0.768
```

A.6.3 Speech Modality

Before Repair:

```
> test <- glm(speech~ConRepair+RobotXP+Sex, data=M.RepairPre,family="binomial")
> summary(test)

Call:
glm(formula = speech ~ ConRepair + RobotXP + Sex, family = "binomial",
data = M.RepairPre)

Deviance Residuals:
Min       1Q   Median       3Q      Max
-0.9848  -0.7606  -0.7012  -0.5915   1.8800

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)     -1.52150    0.45967  -3.310 0.000933 ***
ConRepairRepair -0.37695    0.44121  -0.854 0.392900
RobotXP          0.24382    0.22928   1.063 0.287583
SexMale          0.07469    0.48851   0.153 0.878484
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 163.66  on 146  degrees of freedom
```

```
Residual deviance: 161.24  on 143  degrees of freedom
(2 observations deleted due to missingness)
AIC: 169.24

Number of Fisher Scoring iterations: 4

> predict(test, data.frame(ConRepair = "Repair"), type = "response")
1
0.2
> predict(test, data.frame(ConRepair = "NoRepair"), type = "response")
1
0.2647059
```

After repair:

```
> test <- glm(speech~ConRepair+Sex+RobotXP, data=M.RepairPost,family="binomial")
> summary(test)

Call:
glm(formula = speech ~ ConRepair + Sex + RobotXP, family = "binomial",
data = M.RepairPost)

Deviance Residuals:
Min       1Q   Median       3Q      Max
-1.3984  -0.6294  -0.4579  -0.3022   2.1482

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)      -1.2387     0.7392  -1.676   0.0938 .
ConRepairRepair  -1.5453     0.5647  -2.737   0.0062 **
SexMale          -0.8612     0.5906  -1.458   0.1448
RobotXP           0.5815     0.3052   1.906   0.0567 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 105.497  on 112  degrees of freedom
Residual deviance:  92.785  on 109  degrees of freedom
(2 observations deleted due to missingness)
AIC: 100.78

Number of Fisher Scoring iterations: 5

> predict(test, data.frame(ConRepair = "Repair"), type = "response")
1
0.09230769
> predict(test, data.frame(ConRepair = "NoRepair"), type = "response")
1
0.2916667
```

A.6.4 Interaction Format

```
> modFormats <- glm(Hclass~ConRepair+RobotXP+Sex, data=mergedHo, family="binomial")
> summary(modFormats)

Call:
glm(formula = Hclass ~ ConRepair + RobotXP + Sex, family = "binomial",
data = mergedHo)

Deviance Residuals:
Min       1Q   Median       3Q      Max
```

```
-1.1605  -0.8774  -0.5855   1.2602   2.0263

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept)      -0.4312     0.6274  -0.687   0.4919
ConRepairRepair  -1.1596     0.5213  -2.224   0.0261 *
RobotXP           0.2387     0.2973   0.803   0.4221
SexMale          -0.5635     0.5559  -1.014   0.3108
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 102.986  on 84  degrees of freedom
Residual deviance:  96.371  on 81  degrees of freedom
AIC: 104.37

Number of Fisher Scoring iterations: 4

> predict(modFormats, data.frame(ConRepair = "NoRepair"), type = "response")
1
0.4
> predict(modFormats, data.frame(ConRepair = "Repair"), type = "response")
1
0.175
```

B. Source Code

B.1 Wizard Control Module (Chapter 5)

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;

namespace wizardControl
{
  public partial class Form1 : EZ_Builder.UCForms.FormPluginMaster
  {
    int intCount = 0;

    public Form1()
    {
      InitializeComponent();

    }

    private void Form1_Load(object sender, EventArgs e)
    {

    }


    public override EZ_Builder.Config.Sub.PluginV1 GetConfiguration()
    {

      return base.GetConfiguration();
    }


    public override void SetConfiguration(EZ_Builder.Config.Sub.PluginV1 cf
      )
    {

      base.SetConfiguration(cf);
    }
```

```
    public override void SendCommand(string windowCommand, params string[]
        values)
    {
      if (windowCommand.Equals("Intervention", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetBackColor(btnInter, Color.Red);

      }

      else if (windowCommand.Equals("WaterIntake", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetBackColor(lblWatIn, Color.Red);
        EZ_Builder.Invokers.SetBackColor(waterIntake, Color.Red);
      }
      else if (windowCommand.Equals("WaterGlass", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetBackColor(lblGlas, Color.Red);
        EZ_Builder.Invokers.SetBackColor(ddGlas, Color.Red);
      }
      else if (windowCommand.Equals("yesno", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetBackColor(btnYes, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnNo, Color.Red);
      }
      else
      {
        base.SendCommand(windowCommand, values);
      }

    }


    public override object[] GetSupportedControlCommands()
    {
      List<string> cmds = new List<string>();

      cmds.Add("Intervention");
      cmds.Add("WaterIntake");
      cmds.Add("WaterGlass");
      cmds.Add("yesno");

      //cmds.Add("Intervention");

      return cmds.ToArray();

      //return base.GetSupportedControlCommands();
    }
```

```
    private void btnSetup_Click(object sender, EventArgs e)
    {
      if(txtPID.Text != "")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$PID", txtPID.
            Text);
      }
      else MessageBox.Show(string.Format("No PID supplied"));


    }

    private void btnInter_Click(object sender, EventArgs e)
    {
      intCount++;
      EZ_Builder.Scripting.VariableManager.SetVariable("$intervention",
          intCount);
      EZ_Builder.Invokers.SetBackColor(btnInter, Color.LightGray);
    }


    private void ddCondition_SelectedIndexChanged(object sender, EventArgs
        e)
    {
      if (ddCondition.Text == "High")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$awareCond", "
            high");
      }

      if (ddCondition.Text == "Low")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$awareCond", "low
            ");
      }
    }

    private void ddCondition2_SelectedIndexChanged(object sender, EventArgs
         e)
    {
      if (ddCondition2.Text == "Incremental")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$conditionInc", "
            Inc");
      }

      if (ddCondition2.Text == "Non-Incremental")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$conditionInc", "
            NoInc");
      }

```

```
    }


    private void waterIntake_SelectedIndexChanged(object sender, EventArgs e)
    {
      if(waterIntake.Text == "High")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$waterintake", "high");
      }

      else if(waterIntake.Text == "Low")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$waterintake", "low");
      }
      EZ_Builder.Invokers.SetBackColor(lblWatIn, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(waterIntake, Color.LightGray);

    }

    private void boxPre_Enter(object sender, EventArgs e)
    {

    }

    private void ddWeather_SelectedIndexChanged(object sender, EventArgs e)
    {
      if (ddWeather.Text == "Great")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "great");
      }

      if (ddWeather.Text == "Bad")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "bad");
      }
    }

    private void ddGlas_SelectedIndexChanged(object sender, EventArgs e)
    {
      if (ddGlas.Text == "Untouched")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$glasWater", "untouched");
      }

      else if (ddGlas.Text == "Full")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$glasWater", "
```

```
            full");
        }
        else if (ddGlas.Text == "Halfempty")
        {
          EZ_Builder.Scripting.VariableManager.SetVariable("$glasWater", "
              halfempty");
        }
        else if (ddGlas.Text == "Empty")
        {
          EZ_Builder.Scripting.VariableManager.SetVariable("$glasWater", "
              empty");
        }
        EZ_Builder.Invokers.SetBackColor(lblGlas, Color.LightGray);
        EZ_Builder.Invokers.SetBackColor(ddGlas, Color.LightGray);

      }

      private void btnYes_Click(object sender, EventArgs e)
      {
        {
          EZ_Builder.Scripting.VariableManager.SetVariable("$yesno", 1);
          EZ_Builder.Invokers.SetBackColor(btnYes, Color.LightGray);
          EZ_Builder.Invokers.SetBackColor(btnNo, Color.LightGray);
        }
      }

      private void btnNo_Click_1(object sender, EventArgs e)
      {
        {
          EZ_Builder.Scripting.VariableManager.SetVariable("$yesno", 0);
          EZ_Builder.Invokers.SetBackColor(btnYes, Color.LightGray);
          EZ_Builder.Invokers.SetBackColor(btnNo, Color.LightGray);
        }

      }
    }
}
```

B.2 Wizard Control Module (Chapter 4)

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;
using System.Speech.Synthesis;

namespace SpeechMarkup
{
```

```
public partial class Form1 : EZ_Builder.UCForms.FormPluginMaster
{
  // EZ_Builder.Scripting.Executor _executor;

  SpeechSynthesizer synth = new SpeechSynthesizer();
  int intCount = 0;

  public Form1()
  {
    InitializeComponent();
    // btnSpeak.Click += new EventHandler(btnSpeak_Click);

    // _executor = new EZ_Builder.Scripting.Executor();
  }


  private void Form1_Load(object sender, EventArgs e)
  {

    // Intercept all unknown functions called from any ez-script globally.
    // If a function is called that doesn't exist in the ez-script library, this event will execute
  }

  private void Form1_FormClosing(object sender, FormClosingEventArgs e)
  {
    // Disconnect from the function event
  }

  public override void SendCommand(string windowCommand, params string[] values)
  {
    if (windowCommand.Equals("Intervention", StringComparison.InvariantCultureIgnoreCase))
    {
      EZ_Builder.Invokers.SetBackColor(bntInterv, Color.Red);
    }
    else if (windowCommand.Equals("Correct", StringComparison.InvariantCultureIgnoreCase))
    {
      EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
      EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
    }
    else if (windowCommand.Equals("Incorrect", StringComparison.InvariantCultureIgnoreCase))
    {
      EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
      EZ_Builder.Invokers.SetText(btnIncorrect, "Incorrect");
    }

    else if (windowCommand.Equals("Agreement", StringComparison.InvariantCultureIgnoreCase))
```

```
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Yes");
          EZ_Builder.Invokers.SetText(btnIncorrect, "No");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        }

        else if (windowCommand.Equals("Social", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Social");
          EZ_Builder.Invokers.SetText(btnIncorrect, "Non-social");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        }
        else if (windowCommand.Equals("Feedback1", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "sticks out");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect3, "other way");
          EZ_Builder.Invokers.SetText(btnIncorrect4, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
        }
        else if (windowCommand.Equals("Feedback2", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
        }
        else if (windowCommand.Equals("Feedback3", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "4 dots on one side");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect3, "say yes");
          EZ_Builder.Invokers.SetText(btnIncorrect4, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
        }
```

```
        else if (windowCommand.Equals("Feedback4", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "on edges");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect3, "say yes");
          EZ_Builder.Invokers.SetText(btnIncorrect4, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
        }
        else if (windowCommand.Equals("Feedback5", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "pieces not parallel");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "edges point inside");
          EZ_Builder.Invokers.SetText(btnIncorrect3, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect4, "other two parts");
          EZ_Builder.Invokers.SetText(btnIncorrect5, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.Red);
        }
        else if (windowCommand.Equals("Feedback6", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "asymmetrical slopes");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "wrong sides");
          EZ_Builder.Invokers.SetText(btnIncorrect3, "repeat");
          EZ_Builder.Invokers.SetText(btnIncorrect4, "say yes");
          EZ_Builder.Invokers.SetText(btnIncorrect5, "wrong undo");
          EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
          EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.Red);
        }

        else if (windowCommand.Equals("Feedback7", StringComparison.
            InvariantCultureIgnoreCase))
        {
          EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
          EZ_Builder.Invokers.SetText(btnIncorrect, "wrong dots");
          EZ_Builder.Invokers.SetText(btnIncorrect2, "repeat");
```

```
        EZ_Builder.Invokers.SetText(btnIncorrect3, "say yes");
        EZ_Builder.Invokers.SetText(btnIncorrect4, "wrong undo");
        EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
      }
      else if (windowCommand.Equals("Feedback8", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
        EZ_Builder.Invokers.SetText(btnIncorrect, "piece sticks out");
        EZ_Builder.Invokers.SetText(btnIncorrect2, "mouth wrong direction")
            ;
        EZ_Builder.Invokers.SetText(btnIncorrect3, "repeat");
        EZ_Builder.Invokers.SetText(btnIncorrect4, "say yes");
        EZ_Builder.Invokers.SetText(btnIncorrect5, "wrong undo");
        EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.Red);
      }
      else if (windowCommand.Equals("Feedback9", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetText(btnCorrect, "Correct");
        EZ_Builder.Invokers.SetText(btnIncorrect, "more than 1 dot");
        EZ_Builder.Invokers.SetText(btnIncorrect2, "feet wrong direction");
        EZ_Builder.Invokers.SetText(btnIncorrect3, "bricks on top");
        EZ_Builder.Invokers.SetText(btnIncorrect4, "repeat");
        EZ_Builder.Invokers.SetText(btnIncorrect5, "wrong undo");
        EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.Red);
      }
      else if (windowCommand.Equals("Feedback9_1", StringComparison.
          InvariantCultureIgnoreCase))
      {
        EZ_Builder.Invokers.SetText(btnCorrect, "frog");
        EZ_Builder.Invokers.SetText(btnIncorrect, "other");
        EZ_Builder.Invokers.SetText(btnIncorrect2, "don't know");
        EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.Red);
        EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.Red);
      }


```

```
      else
      {
        base.SendCommand(windowCommand, values);
      }

    }

    public override object[] GetSupportedControlCommands()
    {
      List<string> cmds = new List<string>();

      cmds.Add("Intervention");
      cmds.Add("Correct");
      cmds.Add("Incorrect");
      cmds.Add("WaterIntake");
      cmds.Add("WaterGlas");
      cmds.Add("Agreement");
      cmds.Add("Social");
      cmds.Add("Feedback1");
      cmds.Add("Feedback2");
      cmds.Add("Feedback3");
      cmds.Add("Feedback4");
      cmds.Add("Feedback5");
      cmds.Add("Feedback6");
      cmds.Add("Feedback7");
      cmds.Add("Feedback8");
      cmds.Add("Feedback9");
      cmds.Add("Feedback9_1");

      return cmds.ToArray();

    }


    private void btnSpeak_Click(object sender, EventArgs e)
    {
      /*
      PromptBuilder pb = new PromptBuilder();
      pb.AppendText(textBox1.Text);
      synth.Speak(pb);*/

    }

    private void btnInterv_Click(object sender, EventArgs e)
    {
      intCount++;
      EZ_Builder.Scripting.VariableManager.SetVariable("$intervention",
          intCount);
      EZ_Builder.Invokers.SetBackColor(bntInterv, Color.LightGray);

    }

```

```
    private void ddWeather_SelectedIndexChanged(object sender, EventArgs e)
    {
      if (ddWeather.Text == "Great")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "weatherGood");
      }

      if (ddWeather.Text == "Bad")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$weather", "weatherBad");
      }
    }

    private void ddCondition_SelectedIndexChanged(object sender, EventArgs e)
    {
      if (ddCondition.Text == "Simple")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$condition", 1);
      }

      if (ddCondition.Text == "Metaphorical")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$condition", 2);
      }
    }


    private void AwareCondBox_SelectedIndexChanged(object sender, EventArgs e)
    {
      if (AwareCondBox.Text == "Aware")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$awareCond", 1);
      }

      if (AwareCondBox.Text == "NotAware")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$awareCond", 2);
      }
    }

    private void btnStart_Click(object sender, EventArgs e)
    {
      if (txtPID.Text != "")
      {
        EZ_Builder.Scripting.VariableManager.SetVariable("$PID", txtPID.Text);
      }
      else MessageBox.Show(string.Format("No PID supplied"));

```

```
      if (ddWeather.Text == "")
      {
        MessageBox.Show(string.Format("No weather supplied"));
      }

      if (ddCondition.Text == "")
      {
        MessageBox.Show(string.Format("No condition supplied"));
      }

      EZ_Builder.Scripting.VariableManager.SetVariable("$intervention", 1);
      EZ_Builder.Scripting.VariableManager.SetVariable("$correct", 1);
      EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect", 1);
    }

    private void btnCorrect_Click(object sender, EventArgs e)
    {
      intCount++;
      EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
      EZ_Builder.Scripting.VariableManager.SetVariable("$correct", 1);
      EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
      EZ_Builder.Invokers.SetText(btnCorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect5, " ");



    }

    private void btnIncorrect_Click(object sender, EventArgs e)
    {
      intCount++;
      EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
      EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect", 1);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
      EZ_Builder.Invokers.SetText(btnCorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
```

```
    EZ_Builder.Invokers.SetText(btnIncorrect5, " ");

  }

  private void btnIncorrect2_Click(object sender, EventArgs e)
  {
    intCount++;
    EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
    EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect2", 1);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
    EZ_Builder.Invokers.SetText(btnCorrect, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect5, " ");

  }

  private void btnIncorrect3_Click(object sender, EventArgs e)
  {
    intCount++;
    EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
    EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect3", 1);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
    EZ_Builder.Invokers.SetText(btnCorrect, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
    EZ_Builder.Invokers.SetText(btnIncorrect5, " ");
  }

  private void btnIncorrect4_Click(object sender, EventArgs e)
  {
    intCount++;
    EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
    EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect4", 1);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
    EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
```

```
      EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
      EZ_Builder.Invokers.SetText(btnCorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect5, " ");

    }

    private void btnIncorrect5_Click(object sender, EventArgs e)
    {
      intCount++;
      EZ_Builder.Scripting.VariableManager.SetVariable("$cor", intCount);
      EZ_Builder.Scripting.VariableManager.SetVariable("$incorrect5", 1);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnCorrect, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect2, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect3, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect4, Color.LightGray);
      EZ_Builder.Invokers.SetBackColor(btnIncorrect5, Color.LightGray);
      EZ_Builder.Invokers.SetText(btnCorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect2, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect3, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect4, " ");
      EZ_Builder.Invokers.SetText(btnIncorrect5, " ");

    }
  }


}
```

C. Additional Documents

C.1 Word Order Construction in English

A clause in English (and other languages as well) is the key unit of syntax, capable of occurring independently. It is useful to think of the clause as a unit that can stand alone as an expression of a 'complete'thought'- that is, a complete description of an event or state of affairs.

> John$_{[s]}$ loves$_{[v]}$ Mary$_{[o]}$.
> (Example 1)

> John$_{[s]}$ gave$_{[v]}$ Mary$_{[o]}$ flowers$_{[o]}$.
> (Example 1a)

The basic word order of English is subject-verb-object (SVO) as in the Example 1 and 1a above. The term 'word order' is used to refer to the order of elements in a clause: subject (s), verb (v), object (o), predicate (p), and adverbial (a) and in which order they occur. This basic word order is very similar to the word order found in Danish and German, which generally follow similar rules:

> John$_{[s]}$ elsker$_{[v]}$ Mary$_{[o]}$.
> (Example 2)

> John$_{[s]}$ liebt$_{[v]}$ Mary$_{[o]}$.
> (Example 2a)

Verb phrase (v):
The verb phrase is the central element of the clause, because it expresses the action or state to which other elements relate, and it controls the other kinds of elements and meanings that can be in the clause. In Example 1 above, 'loves' is the verb phrase.

Subject (s):
The subject denotes the most important participant in the action or state denoted by the verb. The subject is a noun phrase and occurs with all types of verbs. In Example 1 above, 'John' is the subject'.

Object (o):
An object is noun phrase which usually follows the verb phrase. Its most common role is the denote the entity affected by the action or process of the verb. We distinguish between direct and indirect object. Direct objects often is the 'doer' or agent of the action, while indirect objects usually denote people receiving something or benefiting from the action of the verb. In Example 1 above, 'Mary' is the direct object, whereas in Example 2 'Mary' is the indirect object and 'flowers' is the direct object.

Predicative (p):
A predicative can be an adjective phrase, a noun phrase, or a prepositional phrase. It follows the verb phrase and (if one is present) the direct object. A predicate characterizes the preceding noun phrase. With predicates we can distinguish between subject predicates, which specify subjects (as in Example 3), and object predicates, which specify direct objects (as in Example 3a).

> Mary$_{[s]}$ was$_{[v]}$ very beautiful$_{[p]}$
> (Example 3)

John[s] found[v] himself[o] in love[p]
(Example 3a)

Adverbial phrase (a):
Some verbs take an adverbial in order to complete their meaning. This is known as an obligatory adverbial. Obligatory adverbials usually express place or direction, although they can also express time or meanings. Optional adverbials can also be added to a clause to add additional information. These are quite flexible and can be added in initial, medial or final positions (see Examples 4,4a,4b)

Today[a] John[s] met[v] Mary[o].
(Example 4)

John[s] often[a] meets[v] Mary[o].
(Example 4a)

John[s] met[v] Mary[o] today[a].
(Example 4a)

While the Danish and German languages exercise a similar kind of flexibility with the position of adverbials it can have an effect on the word order. Specifically, the subject-verb order inverts in Danish and German when placing adverbials in the initial position. The sentence from Example 4 is in Example 5 and 5a translated to Danish and German, respectively.

I dag[a] mødte[v] John[s] Mary[o].
(Example 5)

Heute[a] hat[v] John[s] Mary[o] getroffen[v].
(Example 2a)

Danish and German learners of English should thus take extra care when positioning adverbials in a clause.

C.2 Task sheet for Chapter 3

Welcome to the experiment. We are currently developing a robot that can help people improve their English. To do that, we need people to train its vocabulary and its ability to construct sentences in English, and for this we need your help.

Your Task:

Task 1:
Read the leaflet on sentence structure in English.

Task 2:

Explain in your own words (but in English) to the robot what the functions of these word classes fulfill:
- Subject
- Verb phrase
- Object
- Predicative
- Adverbial

Try to do this without using the instruction sheet.

Now you need to use the letter blocks on the table to construct different sentences in English. When you hold up one block in front of the robot, it will be able to read the word out loud. Once you've finished your sentence on the table the robot is per your construction also able to read out the entire sentence.

When you pick up nouns describe the attributes (size, shape, color, taste, use, etc.) of the specific noun to the robot before you proceed with your instruction.

Task 3:

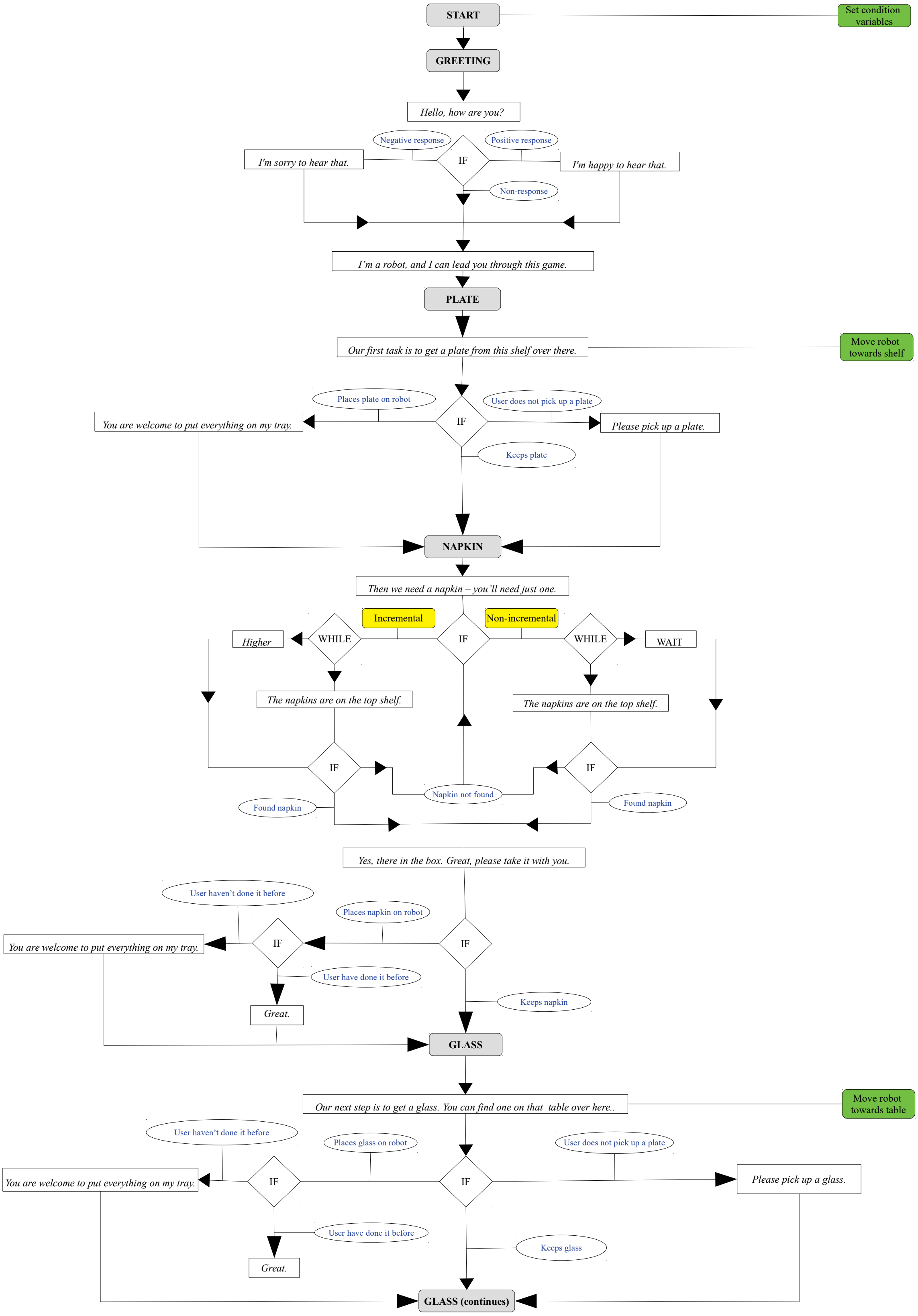
Explain to the robot how to construct a sentence in English using a subject, a verb phrase and an object. Construct at least one sentence with the letter blocks during your instruction.

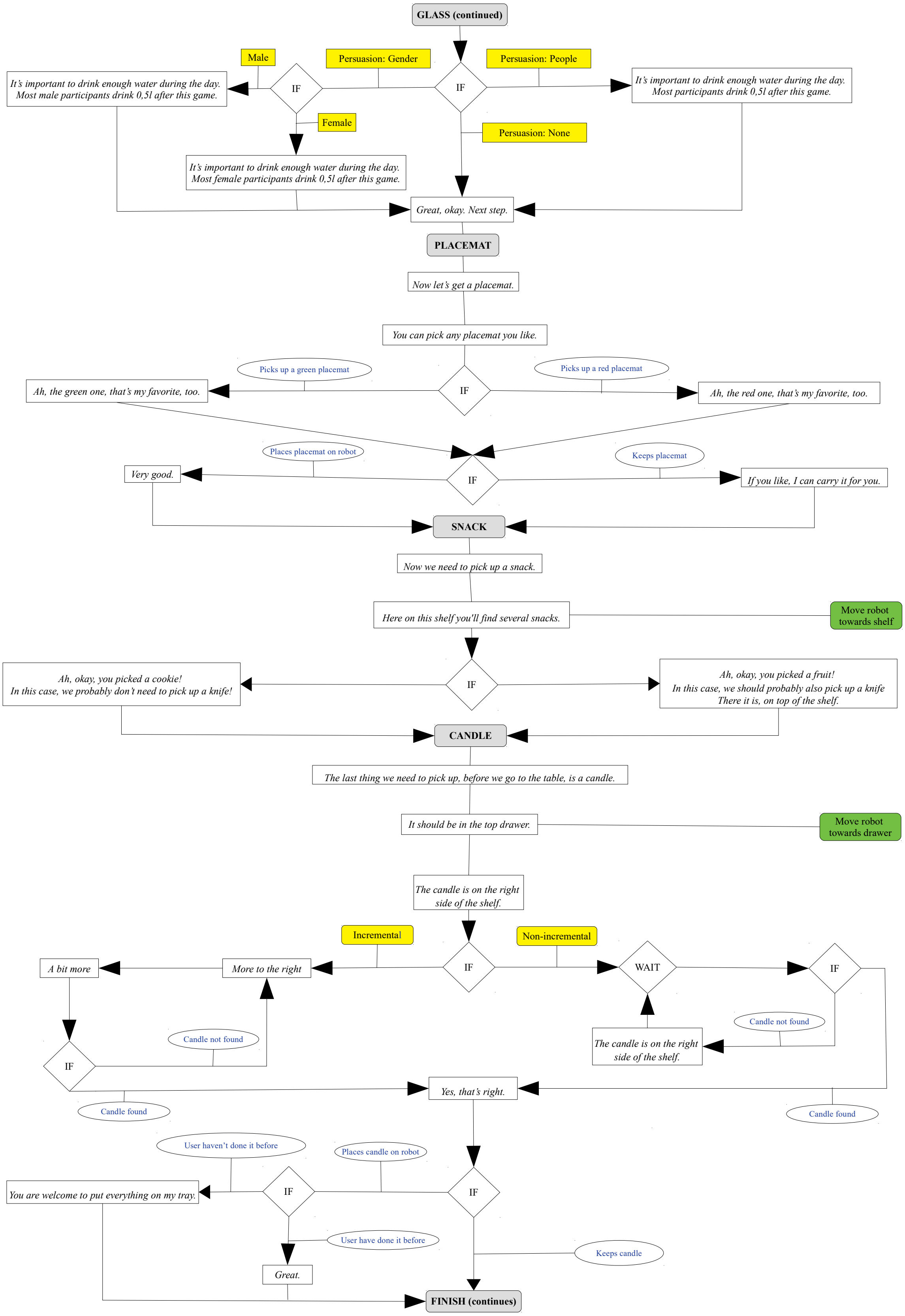
Task 4:

Explain to the robot how to construct a sentence in English using a subject, a verb phrase, an object and a predicative. Construct at least one sentence with the letter blocks during your instruction.

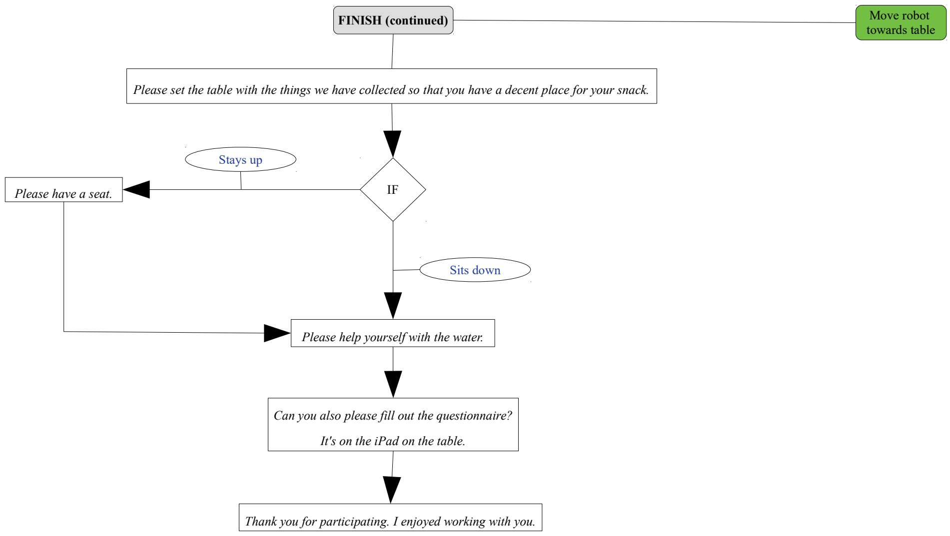
Task 5:
Explain to the robot how to construct a sentence in English using a subject, a verb phrase, an object and an adverbial in initial position. Construct at least one sentence with the letter blocks during your instruction.

C.3 Interaction Protocol for Chapter 6





GLASS (continued)
Male
Persuasion: Gender
Persuasion: People
It's important to drink enough water during the day.
Most male participants drink 0,5l after this game.
IF
IF
It's important to drink enough water during the day.
Most participants drink 0,5l after this game.
Female
Persuasion: None
It's important to drink enough water during the day.
Most female participants drink 0,5l after this game.
Great, okay. Next step.
PLACEMAT
Now let's get a placemat.
You can pick any placemat you like.
Picks up a green placemat
Picks up a red placemat
Ah, the green one, that's my favorite, too.
IF
Ah, the red one, that's my favorite, too.
Places placemat on robot
Keeps placemat
Very good.
IF
If you like, I can carry it for you.
SNACK
Now we need to pick up a snack.
Here on this shelf you'll find several snacks.
Move robot
towards shelf
Ah, okay, you picked a cookie!
In this case, we probably don't need to pick up a knife!
IF
Ah, okay, you picked a fruit!
In this case, we should probably also pick up a knife
There it is, on top of the shelf.
CANDLE
The last thing we need to pick up, before we go to the table, is a candle.
It should be in the top drawer.
Move robot
towards drawer
The candle is on the right
side of the shelf.
Incremental
Non-incremental
A bit more
More to the right
IF
WAIT
IF
Candle not found
Candle not found
IF
The candle is on the right
side of the shelf.
Yes, that's right.
Candle found
Candle found
User haven't done it before
Places candle on robot
You are welcome to put everything on my tray.
IF
IF
User have done it before
Keeps candle
Great.
FINISH (continues)


FINISH (continued)
Move robot towards table
Please set the table with the things we have collected so that you have a decent place for your snack.
IF
Stays up
Please have a seat.
Sits down
Please help yourself with the water.
Can you also please fill out the questionnaire?
It's on the iPad on the table.
Thank you for participating. I enjoyed working with you.