Approaching Human-Like Spatial Awareness in Social Robotics

An Investigation of Spatial Interaction Strategies with a Receptionist Robot

Patrick Holthaus

Declaration of Authorship

According to Bielefeld University's doctoral degree regulations §8(1)g:
I hereby declare to acknowledge the current doctoral degree regulations of the Faculty of Technology at Bielefeld University. Furthermore, I certify that this thesis has been composed by me and is based on my own work, unless stated otherwise. Third parties have neither directly nor indirectly received any monetary advantages in relation to mediation advises or activities regarding the content of this thesis. Also, no other person's work has been used without due acknowledgment. All references and verbatim extracts have been quoted, and all sources of information, including graphs and data sets, have been specifically acknowledged. This thesis or parts of it have neither been submitted for any other degree at this university nor elsewhere.

Patrick Holthaus

Place, Date

Approaching Human-Like Spatial Awareness in Social Robotics

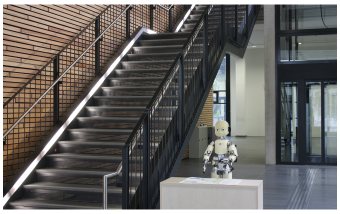
An Investigation of Spatial Interaction Strategies with a Receptionist Robot

Patrick Holthaus

December 2014

A doctoral thesis presented for the degree of
Doctor of Engineering (Dr.-Ing.) at

Faculty of Technology
Bielefeld University
Applied Informatics
Inspiration 1
33619 Bielefeld
Germany



Reviewers

Dr.-Ing. habil. Sven Wachsmuth
Prof. Dr. Giorgo Metta

Examination Board

Prof. Dr.-Ing. Stefan Kopp
Dr. Christina Unger

Defended and approved at Tuesday 10^{th} February, 2015

Printed on permanent paper as per ISO 9706.

Acknowledgments

For making it possible to complete my thesis, I owe special thanks to a lot of people. First of all, I would not have had the possibility to even study without the overwhelming support of my parents and family who always believe in my abilities and supported me not only financially but also with their encouragement in numerous other ways. My lovely wife Natalie however, is the one who deserves most credit because she provides me with so much optimism and confidence, paired with the knowledge that she always backs me up despite the fact that developing a doctoral thesis sometimes is a very time consuming and resource demanding endeavor. Thank you for being there for me in every way imaginable.

Besides that, especially my social environment really backs me up. I truly appreciate that there are still very active friendships with people from my home area as well as those originating in the years of study or as a research assistant. Thank you for diverting my thoughts from work and reminding me of recreational activities albeit also receiving far too little attention from my side.

Thank you Gerhard, Franz, and Britta for granting me the possibility to work with the applied informatics group which has always been very communicative in conjunction with a socially active and friendly atmosphere. From my colleagues, particularly Thorsten and Ingo have helped a lot in introducing me to the group and their topics in the beginning. Also to the people I have shared offices with, it was a pleasure as well as a personal and professional gain spending the time with you.

My thank also goes to Sven, who has been a great supervisor contributing valuable input during the identification stage and progress of this work. Furthermore, I would like to thank Giorgo Metta to also instantly agree on reviewing my thesis. Finally, I very much value the very spontaneous consent of Torben and Peter to proofread my thesis. Thank you.

Abstract

This doctoral thesis investigates the influence of social signals in the spatial domain that aim to raise a robot's awareness towards its human interlocutor. A concept of spatial awareness thereby extends the robot's possibilities for expressing its knowledge about the situation as well as its own capabilities. As a result, especially untrained users can build up more appropriate expectations about the current situation which supposedly leads to a minimization of misunderstandings and thereby an enhancement of user experience.

On the background of research that investigates communication among humans, relations are drawn in order to utilize gained insights for developing a robot that is capable of acting socially intelligent with regard to human-like treatment of spatial configurations and signals. In a study-driven approach, an integrated concept of spatial awareness is therefore proposed. An important aspect of that concept, which is founded in its spatial extent, lies in its aspiration to cover a holistic encounter between human and robot with the goal to improve user experience from the first sight until the end of reciprocal awareness. It describes how spatial configurations and signals can be perceived and interpreted in a social robot. Furthermore, it also presents signals and behavioral properties for such a robot that target at influencing said configurations and enhancing robot verbosity.

In order to approve the concept's validity in realistic settings, an interactive scenario is presented in the form of a receptionist robot to which it is applied. In the context of this setup, a comprehensive user study is conducted that verifies the implementation of spatial awareness to be beneficial for an interaction with humans that are naïve to the subject. Furthermore, the importance of addressing an entire encounter in human-robot interaction is confirmed as well as a strong interdependency of a robot's social signals among each other.

Contents

List of Illustrations

Figures

Tables

List of Definitions

Alignment	is a (linguistic) concept describing an automatic adaption process between interlocutors. 9, 14, 15, 23, 24, 28, 33, 45, 56, 113
F-Formation	refers to a formal notation of the mutual orientation of two or more people during a conversation. 21, 22, 25, 31, 52, 66, 81, 84, 108
HRI	Human-Robot Interaction 8–10, 13, 15, 20, 24, 33, 46, 47, 50, 80, 90, 113
Interaction space	is defined as the space resulting from overlapping the peripersonal spaces of two individuals. 23–25, 27, 31, 33, 46, 52, 54, 63, 84, 102, 108
Motioning	is multiple abrupt communicative whole body movements interrupted by short pauses. 24
Peripersonal space	corresponds to a persons reaching space, directly around one's body. 22, 25, 33, 45
Proxemics	denotes research about human's perception and usage of space. 16, 25, 31, 52, 54, 66, 67, 108, 116

SFP orientation

sociofugal-sociopetal orientation 17, 21

Signal

describes any act or structure to alter another individuals behavior. 13, 16, 24

Social robot

is a robot that is capable of exchanging social signals with a human. 3, 7, 11, 16, 23, 25, 27, 29, 33, 45, 51, 113

Spatial prompting

is the usage of active strategies aimed to influence another's spatial configuration. 24, 25, 31, 33, 52, 67, 84, 86, 91, 97, 102, 108, 116

Spatial signal

is a signal that communicates social aspects of spatial configurations between interaction partners. 16, 17, 24, 28, 29, 32, 47, 51, 52, 87, 90, 110, 113

List of Propositions

1. Robots as Social Entities

Today, most humanoid robots such as HRP-4 (Kaneko et al., 2011) and iCub (Metta et al., 2010) are developed primarily for research purposes. Only in very distinct cases, robots are employed in real-world applications like certain forms of therapy for elderly people (cf. Wada & Shibata 2007) or shopping guides (cf. Gross et al. 2009). Nevertheless, some humanoid robots, e.g. NAO by Aldebaran[1], are already commercially available and thus distributed in public. Robotic agents therefore still play a small but steadily increasing role in the daily life of naïve users.

Whereas a great portion of research on humanoids involves rather fundamental functions like walking (cf. Collins et al. 2005), more interactive features such as spoken dialog are also a common topic of robotics researches. Rickert et al. (2007), as an example, show how to develop an integrated dialog system for humanoid robots involving robot vision and motor abilities.

Imagine such a humanoid robot that is equipped with a dialog resides in the entry hall of a university building acting as a receptionist on an information desk (cf. Fig. 1.1). It has knowledge about the ground plan in order to give directions. A freshman or somebody from another institute enters the building and wants to know the way to the auditorium. He walks towards the desk and the following fictional dialog occurs:

H(uman) Hello.

R(obot) Hello, what can I do for you?

H Where is the auditorium?

R You have to go this way. [Robot shows direction]

H Thank you.

---

[1] http://aldebaran.com/en/humanoid-robot/nao-robot (visited: Wed 29th Apr, 2015)

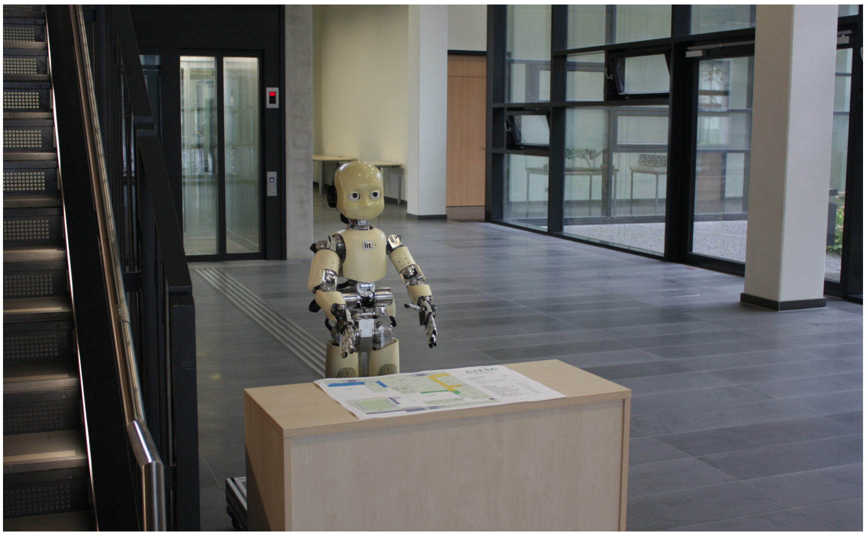


Figure 1.1.: A receptionist robot behind a desk in the entry hall of a university building. The map in front of the robot delineates the building plan.

While such a dialog seems to reasonably match the situation, even in this small example there are a lot of possible pitfalls and misunderstandings. According to Stubbs et al. (2007), particularly highly autonomous robot systems can be difficult to comprehend for a human if they only exhibit a low amount of transparency regarding their decision making. For example, in the very beginning of the dialog, it might not be clear to the human that the robot is actually able to help with the way-finding problem when he enters the room. As a consequence, an uninformed user could as well pass the desk and the robot so that the dialog would never happen. Additionally, it is completely unknown whether the robot is actually capable of understanding speech, so it is somewhat unlikely that the human addresses the robot verbally or even articulates a greeting.

As a possible solution, Breazeal et al. (2005) find an efficient way to rise the robustness of an interaction between human and robot in providing the robot with social functions. Such (mostly) nonverbal signals (cf. Chap. 2.1.2) are intended to reveal the robot's current state to a user so that possible confusions are being minimized.

According to Argyle (1969, Pg. 127), humans as well emit a variety of social signals during interactions that can be used by the other to infer information about themselves, e.g. personality, emotions, and attitude. Social behaviors in humans are on the one hand based on innate factors and on the other hand "embodied in cultural rules and norms" (Argyle 1969, Pg. 26). They are driven by factors such as dependency, affiliation, and achievement motivation.

Although such social skills, for example empathy (cf. Tapus & Mataric 2009), can be emulated, robots as opposed to humans are not intrinsically social. In order to build a social robot, Hegel et al. (2009) state that it is a necessary prerequisite to incorporate physical cues that foster social interaction. The more a robot is equipped with human-like features, Krach et al. (2008) find as neurological evidence, the more humans treat the robot as if they had a mind. Riether et al. (2012) in addition are able to demonstrate the effect of social facilitation (cf. Triplett 1898), i.e. an impact of robot presence on task-solving performance.

Nevertheless, for an interactional purpose it is not sufficient to only communicate the capabilities and inner state of a robot but also to perceive human behaviors and interpret them correctly (cf. Adams et al. 2000). On the basis of this knowledge, the robot is then able to emit appropriate social behavior itself. For example, Lee et al. (2010) present a receptionist robot that is treated as a social being as opposed to an information kiosk.

In summary, it can be stated that an interaction between robot and human can be enhanced in terms of robustness and usability if the robot is able to correctly interpret and exhibit social behaviors. This requires a certain level of mutual understanding in both agents. According to Sheridan (1997), one of the early but still prevailing challenges in the creation of such a *social robot* is therefore to provide humans and robots with appropriate models about their counterpart.

Besides robustness and usability, further aspects of the interaction can be influenced positively with the help of social features. For example, Lee et al. (2006) demonstrate an impact on the level of perceived personality which plays an increasingly important role if one considers the rising impact of robots in society. Furthermore, Kanda et al. (2004) observe in a field study that children lose interest in the robot after a longer interaction because the robot lacks a technical as well as a social challenge for their partners. Additionally, Leite et al. (2013) elaborate on a variety of long-term studies with social robots that reveal benefits in interaction scenarios. Social intelligence therefore enriches the user experiences and increases the long-term viability of humanoid robots in different ways. As a consequence, many researchers (cf. Fong et al. 2003a) create their robots in a socially intelligent manner.

If one now recalls the original dialog between the receptionist robot and the visitor, the interaction could be enhanced in many different ways with the help of social signals. The robot could for example expose its detection of the person entering by raising its head. The human is then aware of the fact that the robot is switched on could conclude that it is available for an interaction. Pitsch et al. (2011), for example, demonstrate such an approach in a museum guide robot which can pro-actively communicate its availability by constantly looking for an interaction partner. Following a similar strategy, Lütkebohle et al. (2009a) use a dialog system that provides a robot with the ability to formulate sentences on its own based on saliency and dialog history to make the robot's capabilities more transparent to the user.

Nevertheless, these works only consider single aspects or stages of an interaction, like capturing attention beforehand or maintaining interest during dialog. This thesis, in contrast, multi-modally covers social aspects in a more complete course of action. It addresses a comprehensive encounter between one human and a robot involving the person entering a room and narrowing the robot, a common dialog that includes mutual greeting and farewell statements, and a departure phase in the end.

Humans permanently communicate with the help of subtle (subconscious) hints even when they are still far apart from each other. Already the positioning gives some indication of possible intentions (cf. Chap. 2.2). Such spatial hints therefore qualify especially for researching during the period of an entire scenario. As a consequence, this work specializes on spatial awareness by developing strategies that enable a humanoid robot to on the one hand interpret spatial communication of a human and on the other hand exhibit such signals itself.

To thoroughly investigate social (including spatial) relationships between humans and social robots, a common practice is to "create real [robot] systems and then evaluate these systems using experiments with human subjects" (Goodrich & Schultz 2007, Chap. 6.1). The presented writing therefore describes the development, implementation, and evaluation of spatial aspects of social competence.

In the next chapter, necessary theoretical concepts for this thesis will be outlined. Afterwards, Chapter 3 proposes an integrated concept of spatial awareness as an approach for fostering mutual understanding between human and robot. The receptionist robot and its basic course of action is described in Chapter 4 as a possible application in which the viability of proposed spatial awareness can be evaluated. Chapter 5 consequently relates the concept to the scenario. It thereby describes in greater detail which strategies have been developed to achieve certain effects and how they are incorporated into the existing setup. The effectiveness of introduced spatial awareness is subsequently evaluated in Chapter 6. Eventually, the thesis is concluded in Chapter 7 with a reflection of its contribution and impact as well as a short summary.

2. Communication among Humans and Robots

Research on human social behavior provides insights on several conscious and unconscious communicative habits of humans. To enhance interactions that involve robots by a social component, one has to find out whether such (or similar) behaviors can also be observed towards them. Inferences have to be drawn on how these insights can then be applied to a *social robot* in order to be beneficial for the interaction. Based on such an approach, this chapter introduces related concepts of communication between humans and applies them into interactions between a human and its robotic interlocutor.

To precisely categorize the thesis' intended topic, this chapter will give an introduction to related research and put this work into context. It begins with a broader view on the subject, explaining the underlying concepts of communication. Afterwards, closer focus will be laid on individual phenomena that are directly affecting the implemented strategies on the robot. Therefore, more generic aspects of communication between humans and robots will be introduced firstly in Section 2.1. Secondly, as this thesis examines the topic of spatial communication, sharper focus will be laid on the social properties of shared spaces in Section 2.2. The chapter will be closed with a short review of the discussed phenomena in Section 2.3.

2.1. Prerequisites for Human-Robot Interaction

To successfully implement a *social robot* into society, current robotics research heavily relies on social sciences and the results from experiments with only human participants. Such insights are commonly believed to be beneficial for a robotic system with an interactional purpose (Restivo, 2001; Breazeal et al., 2008). Thereby, one has to consider that communication between robots and humans is asymmetrical due to the fact that robots do not have the same cognitive capabilities as humans (cf. Nomikou et al., 2013; Fischer et al., 2014). Accordingly, Dautenhahn (2007) states that concepts from interactions between humans cannot be transfered directly to *Human-Robot Interaction (HRI)*. Wrede et al. (2010) therefore suggest that especially the robot's feedback signals have to be designed diligently with regard to their social impact. Also, findings have to be validated and tested extensively on the robot according to Kahn et al. (2010) in order to be fully established as an efficient system.

Human spoken dialog is usually accompanied by gestures which essentially cover similar kinds of information (cf. Kendon 1997). At the same time, interactions in *HRI* are similarly composed of verbal and non-verbal communicative acts, as Mavridis (2014) recently describes in a comprehensive overview. Expressive gestures in conjunction with a dialog system therefore are commonly believed to help a social robot to convey information more precisely. Dialog systems of artificial agents can for example be enhanced with co-verbal gestural utterances as for example outlined by Kopp & Wachsmuth (2004) and implemented by Salem et al. (2013a). As a complementary approach, this work does neither directly address iconic or lexical gestures during dialog nor dialog itself. Instead it lays its focus to supportive strategies that are meant to accompany dialog and gesture.

In the next parts of this chapter, different underlying theories from social sciences that are a prior to this thesis are being introduced. At first, it will be investigated how mutual mental models between humans are discussed as linguistic phenomena and point out how they can be applied to human and robot interactions in Section 2.1.1. Afterwards, social signals are introduced in Section 2.1.2 as a carrier of information to build up such models. Finally, in Section 2.1.3 focus will be laid on the deviating expectations humans have towards a robotic interaction partner as opposed to a human counterpart.

2.1.1. Mutual Understanding

In order to communicate successfully, humans have to establish a certain degree of mutual understanding. According to Clark & Wilkes-Gibbs (1986), it is a crucial criterion for all dialog partners to believe that the receiver(s) sufficiently understand the meaning of a contribution. They describe the process of grounding as the mutual establishment of mental models that contain such information via negotiation. The resulting so-called common ground allows for shared knowledge, beliefs, and assumptions and thus is an essential prerequisite for effective communication.

Moreover, also in communication between humans and robots, mental models are believed to help understand one another. For example, Sheridan (1997) names the provision of humans with models of machines and vice versa as one of the central challenges in social robotics. A major effort therein consists in mapping the other's inner state to an understandable representation. Grounding as a principle therefore has been effectively incorporated in a number of artificial dialog systems for humanoid robots (eg. Eimler et al. 2010). In addition, Peltason et al. (2013) argue that although preconditions vary greatly in comparison to human-human interactions, similar grounding effects can be observed in *HRI* as well.

In contrast, Pickering & Garrod (2004, Pg. 188) introduce interactive *alignment* as a linguistic adaption "process [that] is automatic and only depends on simple priming mechanisms". The alignment model describes a resource-efficient and implicit construction of common ground. By the assumption that priming activates similar patterns in the partner, an explicit model of beliefs about the other's understanding is only needed as a fall-back mechanism. Due to the propagation of alignment through the interconnected linguistic representations (Fig. 2.1), it allows for successful communication solely with the help of priming. The authors claim alignment to occur on different levels of mental representations: Beginning from phonetic and lexical representations, e.g. usage of the same linguistic construct, over semantical alignment up to the interlocutors situation models (cf. Zwaan & Radvansky 1998). The interactive alignment model also has inspired further research, supporting its general ideas. For example, Pardo (2006) finds evidence for phonetic alignment and Richardson et al. (2007) discuss the interplay between eye-movement coupling and dialog. Menenti et al. (2012) furthermore examine a neural basis for alignment.

Despite that, the universality of the alignment concept is also controversially discussed. Already the open peer commentary regarding the article of Pickering & Garrod (2004) contains arguments by Brown-Schmidt & Tanenhaus; Schiller & de Ruiter against priming as the only mechanism behind alignment. Krauss & Pardo consequently argue to allow for more conscious decisions other than priming. In addition, not all aspects of dialog can be explained by alignment phenomena. Carbary et al. (2010) for example could not find any correlation between syntactic alignment and joint task improvement, indicating that some form of planning has to play a role be-

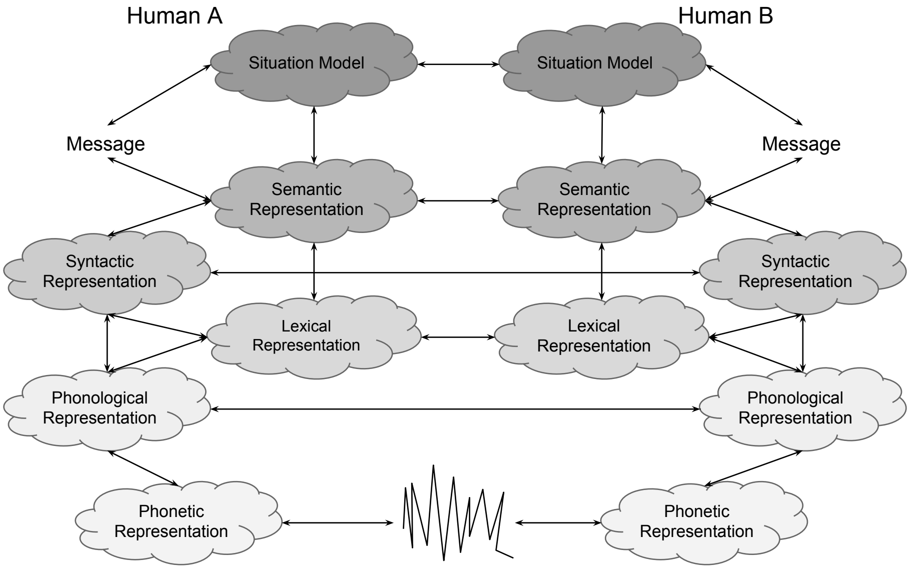


Figure 2.1.: The alignment model as described by Pickering & Garrod (2004, Pg. 177). Interlocutors A and B align their mental representations in a dialog across different linguistic levels including the situation model.

sides automatic alignment. Induced by discussions in the dialog and robotics communities, a further exploration of alignment phenomena as the common foundation for models in *HRI* has inspired many interdisciplinary projects in the collaborative research cluster "Alignment in Communication" at Bielefeld University in which this thesis is embedded in. Wachsmuth et al. (2013) give an overview of selected topics within the cluster.

Braningan & Pearson (2006) state that alignment effects can be observed between humans and robots and that they are often stronger than those between humans. Nonetheless, human-robot alignment is strongly modulated by the individual's concept of the robot as a communication partner. Thus, the collaboration between linguistics and computer science, as propagated in the cluster, constitutes a mutual benefit: On the one hand, social behaviors of robots can be outlined, implemented, and verified based on the alignment principle. On the other hand, the linguistic notion of alignment can be extended to artificial agents and modalities besides speech.

Among computer scientists in the cluster, the idea of alignment is manifested in many of the cluster's robotic systems. Swadzba et al. (2009) successfully implement a computational model of three-dimensional scene representations to align their mobile robot's representation to a human understanding of spatial structures in rooms. Damm et al. (2012) present a computational model for an emotional alignment between an artificial agent and a human. Both approaches follow the idea of expanding the alignment phenomena from speech to spatial arrangements or emotions.

In the same way, spatial interaction strategies in this thesis (cf. Chap. 5) are guided by the idea of alignment between human and robot. Due to a hierarchical representation of knowledge in a technical system that is based on a human-like structure, a solid basis for mutual understanding is established. With such a design as a foundation for reasoning, the robot's spatial presence can be aligned to an interlocutor's behavior. As a consequence, the robot furthermore is able to provide an idea of its internal understanding of the humans actions and representations. The alignment model therefore provides a possibility to increase a *social robot's* context awareness with regard to spatial configurations and thereby its usefulness in general.

Figure 2.2 describes the manifestation of the alignment principle into a technical system that can align spatially to an interlocutor. The graphic is inspired by the original as depicted in Figure 2.1. Although the internal representations in humans and robots are inherently different, alignment can occur on different levels of representations by persistently communicating the robot's internal states to the human as well as interpreting the humans verbal and non-verbal signals.

2.1.2. Conveying Information

Mental models as a basis for mutual understanding of independent agents can not be established in isolation but only through the exchange of information. Humans in parallel to speech employ nonverbal communication in their daily interactions as presented by Knapp et al. (2013) in a comprehensive overview. According to Patterson (1982), the most basic function of nonverbal behaviors thereby consists of providing each other with information. As spatial communication as well (cf. Chap. 2.2) to a large degree is of non-verbal nature, such signals play a crucial role in extending an interaction to the spatial domain and enabling spatial awareness in humans.

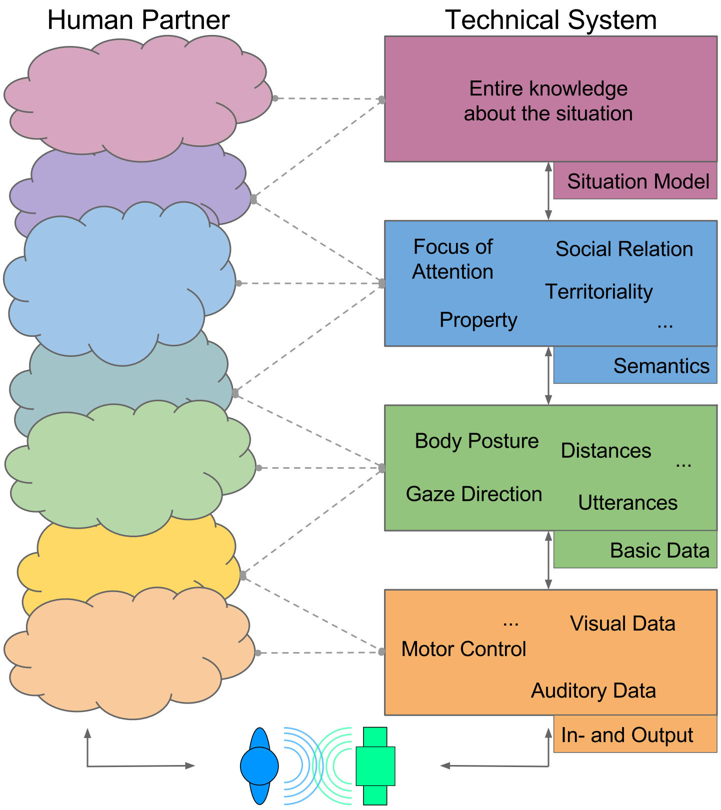


Figure 2.2.: Schematic depiction of an interactive technical system and a human partner. The system consists of exemplary modules that are designed to align with structures that represent similar information in the human counterpart. The most abstract representation in form of the situation model can be found on top, the most specific and low-level representation at the bottom. Structures in general are not restricted to linguistic representations. Structures in the human are depicted as fuzzy clouds in order to indicate that their kind of representation as well as the boundaries between them remain unspecified. Information between human and system can only be exchanged via signals in different modalities and therefore have to be propagated through the system from the bottom.

Accordingly, nonverbal social signals in robots serve a very specific purpose: They facilitate an interaction between human and robot by giving the human subtle hints on how to interact with the robot and therefore support mutual understanding by the creation of mental models. As an example, Pitsch et al. (2013) find that gaze strategies can be used to pro-actively shape the humans behavior during a tutoring situation.

Still, speech certainly is one of the most prominent and most intensively researched means of communication in social robotics (cf. Fong et al. 2003b). Nonetheless, Goodrich & Schultz (2007, Chap. 4.2) specify a considerable proportion of non-verbal messaging between humans and robots. This section therefore elaborates on the concept of a non-verbal signal and relates it to *HRI*. After describing how social signals are affecting interactions with robots, the concept of joint attention is briefly introduced as an important opportunity for information exchange utilized in this thesis.

In the domain of humans (and animals), Maynard-Smith & Harper (2004) have defined an exchange of information from one entity to another as a *signal*. Signals carry information from the sender to the receiver, thereby causing a certain cost to the sender, such as energy or time-involving costs. In contrast to the categorization in Allwood (2002), this thesis refers to the term signal in the passive (e.g. color, appearance) and the active (gesture, speech) case as they all influence the receiver's behavior.

After Hegel et al. (2011), the specific notion of signals and their properties can be beneficial to guide developers in the design of skills for a social robot and the analysis of existing behaviors. They claim the important distinction between artificial signals such as LEDs or acoustic beeps and human-like signals such as a facial expression. While the study of artificial signals certainly is a research topic on its own, especially in human-computer interaction, it is not a part of this thesis. Instead, it focuses on the question whether human-like social signals can be beneficial for the user experience when interacting with a robot.

Such signals have to be designed carefully because they involve a certain degree of deception towards the human which lowers the signal's reliability (cf. Donath 2011, Pg. 4). Robots do not need to blink as humans or maintain eye-contact. Nevertheless, many existing platforms (e.g. Yoshikawa et al. 2006; Yonezawa et al. 2007) implement this type of features in order to appear as social characters. A consistent course of action is thereby crucial in designing a believable character according to Simmons et al. (2011).

At the same time, such signals only work because humans tend to anthropomorphize certain properties of robots (Duffy 2003; Eyssel et al. 2010). According to Epley et al. (2007), especially in humanoid robots this tendency can be fostered by creating social bonds using human-like signals. Humans then behave similar when interacting with a robot as if they are interacting with another person. Users of robotic systems thus benefit because they do not have to learn specific ways to operate a robot but use them intuitively. Developers of robotic systems also capitalize on the fact that humans react to the behaviors of the robot in a more predictable way as they have an extended knowledge of the robot's capabilities.

A very prevalent way to transfer information to an interaction partner is the mechanism of joint attention as summarized in Moore & Dunham (1995). Humans older than one signal each other their current focus of attention with the help of gaze patterns which reduces uncertainty in the dialog. Mundy & Newell (2007) even argue that joint attention drives the development of social cognition in humans. Joint attention also seems to facilitate the creation of *aligned* mental models according to Brown-Schmidt et al. (2005, Pg. 167).

Joint attention has also been adopted by various researchers on their social robots. Wrede et al. (2009) for example, describe joint attention as a key mechanism for learning in an interaction with a humanoid robot. Yu et al. (2010) investigate joint attention in a learning scenario as well. They find that humans better regard a robot if it shows patterns of joint attention as well as unnatural behavior in other conditions. Furthermore, Caraian & Kirchner (2013) argue that joint attention emitted by a robot can take severe influence on object choices similar as in human-human interactions.

2.1.3. Expectations on Social Robots

In order to fulfill the intended purpose, a signal has to be efficient with regard to the task which is currently being solved, i.e. the human has to be able to interpret the signal appropriately (cf. Breazeal 2003). Lohse et al. (2007) find that already the visual appearance alone has significant influence on the purpose attributed to a robot. Riek et al. (2010), for example, discover that the style of a robot's social gesture has a significant influence on the reaction time of the user. They claim that more abrupt gestures require less mental processing and thus lead to a smoother interaction.

Lohse (2011) states that effectiveness to a large degree depends on the expectations a user has of the robot and also the situation. Accordingly, robot signals have to be clear and match the current situation in order to lead a human through the interaction. Paepcke & Takayama (2010) confirm that a manipulation of users in terms of wrong expectations can lead to disappointment about robot capabilities. Accordingly, Breazeal (2004) uses expressive cues to consistently inform the human about the robot's internal state. This way, social signals provide a possibility have to reveal the robot's capabilities if they not raise the expectations too high or point into a wrong direction. As a consequence, a human is able to maintain an updated mental model of the robot, which makes the interaction more intuitive and benefits both partners.

Likewise, Duffy & Joue (2004) state that user expectations play a crucial role for the design of social signals in *HRI* and this is also the case for this thesis. Robot signals in general should be conceptualized with focus on the user so that a common concept of the current contextual information can be developed. Furthermore, by regarding user expectations, a robot is capable to incorporate anticipatory actions into its repertoire of behaviors as stated by Hoffman & Breazeal (2007), which also contributes to a more efficient collaboration.

As an example, spoken dialog is a skill that many humans expect a social robot to be equipped with according to Cappelli et al. (2005). Additionally, Fischer (2011) discover that a carefully designed dialog reduces uncertainty and can guide users during the interaction if it meets the users expectations. Utterances therefore should approximately match the speech recognition's abilities to not confuse the user about the robot's skills. Specific movements that support an utterance, as described by Jung et al. (2013), can also lead to a more pleasurable user experience if they are adjusted to match different phases of a dialog.

Consequently, user expectations play a fundamental role for a well coordinated compilation of interrelated robot strategies presented in Chapter 5: As a first assumption, the suggested behaviors are based on patterns of interactions between humans. If these patterns work as intended for an interaction between human and robot, an *aligned* state between them should emerge. Chapter 6 verifies to what degree such assumptions are valid for the proposed strategies and how they have to be adjusted.

2.2. Spatial Communication with Robots

In this thesis, the influence of spatial awareness on the communication between humans and robots is discussed. The exact term *spatial signal* will be used for acts or properties that fulfill the following criteria: a) It can be considered a *signal*, i.e. have an influence on the mental model of an interaction partner b) The influence changes conceptions related to the spatial configuration of the interaction partners. The signal could either be influenced spatially because it relies on a spatial configuration as an information source (input) or because it changes the spatial configuration itself (output). For example, a spatial signal after this definition would be a repellent reaction to someone who tries to embrace you because the reaction is distance triggered.

In the next sections, research that investigates the presence and usage of spatial signals in interactions will be introduced (cf. Ciolek 1983 for a list of common notations). These concepts lay the theoretical foundations of this thesis as the behaviors introduced in Chapter 5 are directly inspired by them. At first in Section 2.2.1, phenomena that occur on a larger scale, namely interpersonal distances and spatial configurations between persons as a whole, will be discussed. Secondly, Section 2.2.2 introduces the way humans engage in face-to-face situations and how such knowledge can be used for *social robots*. The interpretation and usage of space during the actual conversation is then subject to Section 2.2.3. Finally in Section 2.2.4, the concept of spatial prompting that has recently come up in the social robotics community will be explained.

2.2.1. Interpersonal Configurations

Interpersonal spaces play an often unconscious (cf. Hall 1966, Pg. 115) but nonetheless important role in the daily communication between people. As one of the first researchers, Hall (1966) deeply investigates the topic and gives an extensive and cross-cultural analysis of human spatial behaviors. Consequently, Hall et al. (1968, Pg. 83) established the term *proxemics* as “the study of [hu]man’s perception and use of space”. While proxemics research also covers large scales such as landscape or city designs, this thesis is restricted to so-called dynamic or interpersonal spaces. Also, interaction strategies are based on findings regarding American and European culture because, amongst others, Watson (1972) states that it remains relatively unclear to what extent the cultural background influences spatial behavior.

By the investigation of correlations between occurrences of *spatial signals* and those of distance patterns, proxemics correlates physical distance to social distance between individuals. For experience of spatial relations, i.e. observation and communication (signaling) of social closeness, Hall (1963) defines eight so-called dimensions: postural-sex identifiers, *sociofugal-sociopetal orientation (SFP orientation)*, kinesthetic factors, touch code, retinal combinations, thermal code, olfaction code, and voice loudness. Some of them are not applicable to robots (yet), and therefore only the dimensions that are relevant for this thesis will be introduced briefly:

- *SFP orientation* describes the angular relation of two bodies towards each other. Although in principle this dimension is of linear nature, one can simplify it to eight areas by using categories of 45 degrees.
- Kinesthetic factors determine in which way two persons can reach each other. The closest form would be direct body contact, followed by touching with the elbow or the forearm. Fully extended arms and leaning correspond to the farthest possible values. As there are two persons involved, there are 64 possibilities in total of which 11 suffice to describe everyday situations.
- Retinal combinations categorize where another person appears on ones retina. Persons in focus are sensed in the fovea while someone in the macular or peripheral area is gradually sensed less.

Dynamic space can be divided into four distance classes, dependent on certain boundaries of the eight dimensions, e.g. kinesthetic features. Please confer to Hall et al. (1968, Pg. 92-93) for a graphical depiction of the correlations. Limits of the classes cannot be interpreted as fixed as they obviously vary with body height and arm length. Also, cultural background, personality, gender, and age (Remland et al., 1995) play a role in defining their exact boundaries. Hall (1966, Chap. X) relates the classes to typical occurrences during human dialogs. In the following, the different classes are listed and a short summary of their respective properties is given. Please refer to Figure 2.3 for a schematic depiction of the different classes.

- Intimate distance directly surrounds a persons body and ranges up to 45cm. Others can be detected with almost all senses including haptics and olfaction. The close phase (closer than 15cm) is generally only entered by sexual partners or in special cases like in competitive sports. For most people it is even uncomfortable if strangers enter the far phase. In situations where such close contact is unavoidable (e.g. in elevators), according to Argyle & Dean (1965), people often try to virtually extend the space by fixing their eyes on infinity and reducing their movements to a minimum. In Figure 2.3, limits of the intimate distance are depicted in violet.

- Personal distance spans up to 1.2m. Closer than approximately 75cm, only one person has to extend her arm to reach the other, which defines the close phase. Typically, only close relatives or otherwise very familiar people enter the close personal distance. In the far phase, it is still easily possible to touch each other if both extend their elbows. It is commonly used among good friends and discussing topics with personal involvement. A magenta colored circle surrounds the personal distance in Figure 2.3.

- Social-consultive distance usually occurs in impersonal conversations like business meetings. It reaches about 3.6m, the distance where two people can barely reach each other while stretching their whole body. Also, the mandatory recognition distance ends here. In the closer phase (up to 2.1m), often a greater degree of involvement can be noticed between the interactants. The far phase indicates a more formal character of the dialog. It is limited by a cyan line in Figure 2.3.

- Public distance is only very rarely used in face-to-face interactions. Dialog in the close phase (up to 7.6m) becomes very formal and distant. Everything above that is considered as the far phase and is only used in situations where there is one speaker in front of an audience, e. g. scientific speeches. Figure 2.3 marks the end of the close phase with a dashed blue circle.

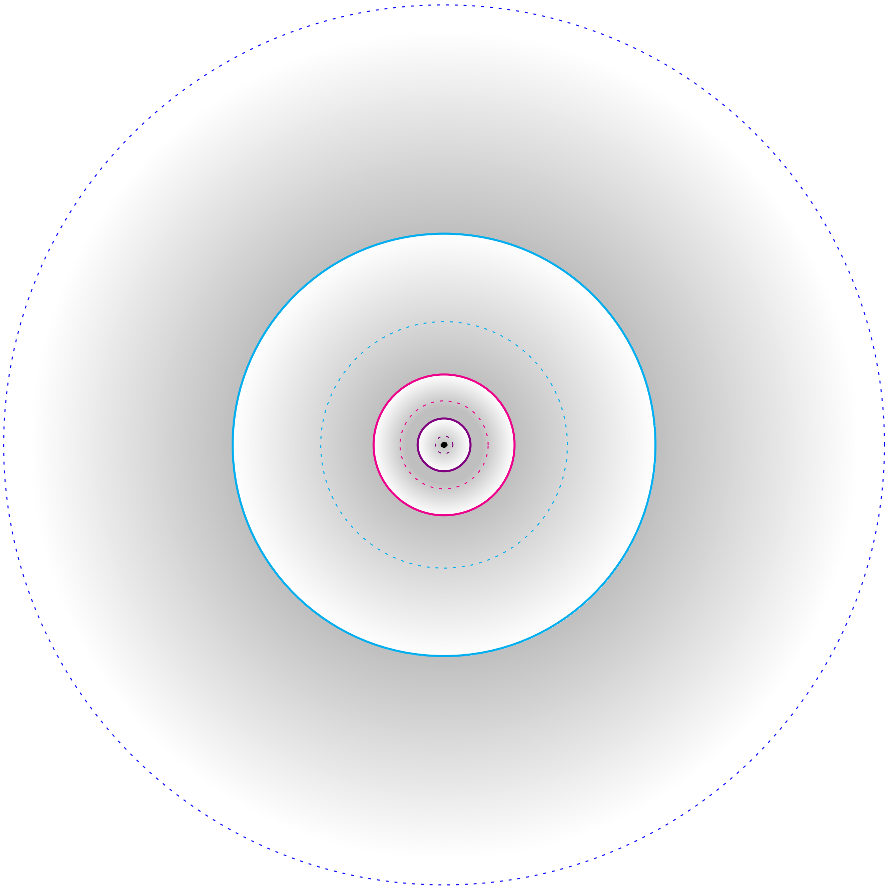


Figure 2.3.: True-to-scale depiction of the distance classes surrounding a person as defined by Hall et al. (1968) in proxemics theory. Solid lines determine the individual borders: violet surrounds the intimate distance (up to 0.46m), magenta marks the personal distance (ranging to 1.22m), and cyan limits the social distance (at 3.70m). Dashed lines indicate the difference between outer and inner phase of the respective class. The dashed line in blue depicts the inner phase of the public distance (7.60m) which itself is unlimited.

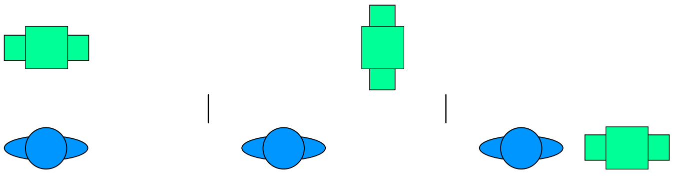
Features of proxemics research seem to be transferable into a computational model without a major effort. Previously, Sundstrom & Altman (1976) have formulated a model of an individuals personal space as a function that maintains an optimal proxemic distance. Such characteristics make it especially interesting for human-computer interaction (cf. Greenberg et al. 2014) as well as for *HRI* (eg. Torta et al. 2011) to implement on devies that are intended to incorporate social functionalities.

Especially the proxemic properties of mobile service robots are of interest in current social robotics science. Kirby et al. (2009), for example, incorporate social conventions directly at a planning level in the navigation module of their robot. Pacchierotti et al. (2006) evaluate the application of a passing-by model based on proxemics. They claim that a carefully designed evasion behavior that respects personal spaces is the most comfortable for naïve users.

Also, Syrdal et al. (2008) investigated what influence the appearance of a robot has on the individuals rating of social distances. In their experiment, participants kept a greater distance towards a more human-like robot as to a more mechanical one. They conclude that this outcome may have resulted due to the participants' expectations of the human-like robot to have some understanding of proximity. Koay et al. (2007) however show that approach distances adjust over time as they are influenced by a habituation effect that occurs in repeated interactions with a robot in a longitudinal study.

Syrdal et al. (2007) as well as van Oosterhout & Visser (2008) and Takayama & Pantofaru (2009) confirm that humans apply distinct spatial models to robots depending on their own body features such as age, height, or gender. According to them, comparable mechanisms as in human-human interactions occur when approaching a robot. Nevertheless, the behavior varies to a certain degree with the experience people have with humanoid robots.

While people certainly do not handle spatial configurations with robots exactly as with humans, many parallels can be observed. Humans apparently seem to attribute some kind of spatial competence to a robot. The findings above suggest that this competence is based on expectations humans have and that their initial approach corresponds to their own experience of social spaces.



(a) Vis-a-vis (b) L-shaped (c) Side-by-side

Figure 2.4.: Sketch of two interactants in exemplary F-formations from above as introduced by Kendon (1990).

2.2.2. Engaging in Conversations

Schegloff (2002, Pg. 329) describes the opening of a conversation between people not as a problem of “who talks first”, but rather the fine-grained decision process of “how the initiation of a conversation is done”. Several context variables such as the relationship between speakers play a role. Kendon (1990, Chap. 6) also states that opening an interaction from afar is a stepwise process which involves engagement and disengagement of mutual gaze at different distances while one person is approaching the other.

At direct encounters however, Albas (1991) additionally finds that people stick to their proxemic habits and even correct interpersonal distances if the other rearranges. In contrast to unfocused interactions, the *SFP orientation* changes with the context of the conversation, especially if there are tasks or objects involved (cf. Ciolek & Kendon 1980). Thus, Kendon (1990, Chap. 7) formalizes the notation of *F-Formations* that humans use to maintain the shared space during an interaction. In order to allow all participants of an interaction direct access to the subject, people for example orient in vis-a-vis, L-shaped, or side-by-side configurations (cf. Fig. 2.4).

This notation is commonly used to structurally analyze interaction situations with regard to their communicative focus (e.g. in Marshall et al. 2011; Gan et al. 2013). For robotics, Yamaoka et al. (2010) present the implementation of a proximity control model for mobile robots based on *F-Formations*. Also, Kuzuoka et al. (2010) apply such a model to structure their interaction and dynamically rearrange visitors in a museum to explain different exhibits.

While proxemics as well as the *F-Formation* system for itself are both subject to a lot of social robotics research, the actual opening of an interaction including the approaching phase has been mostly disregarded. Lütkebohle et al. (2009a), for example, present a robot that is capable of self- or mixed initiative dialog. Nonetheless, the interaction does not start until a person stands directly in front of the table. Others (e.g. Shiomi et al. 2010) use a remote control to initiate an interaction but also ignore the fine-grained entry process and establishment of a common interaction space.

In contrast, Pitsch et al. (2009) open the interaction as soon as the robot detects human activity. It maintains contingent eye contact during the whole approaching phase. Also, Heenan et al. (2013) present a sophisticated system for social interaction opening, but both of them use the relatively small (58cm) NAO robot[1] so it remains unclear how exactly the enormous height difference effects the users behavior (cf. Sect. 2.2.1).

2.2.3. From Reaching Space to Interaction Space

Widely accepted research shows that humans are neurologically equipped with a special representation of their ultimate surroundings, which is called the *peripersonal space* (cf. Holmes & Spence 2004). Most authors agree that the human brain represents it fundamentally different than anything that is farther away. Peripersonal space is also known as reaching space because it covers exactly the area within arm length of an individual.

While the peripersonal space itself describes a characteristic that is only applicable to single individuals, Lloyd (2009) gives an extensive overview of the research regarding the peripersonal space and relates it to interpersonal communication. Not only it is to a certain degree possible for humans to extend the body space by using tools that enhance their reaching capabilities according to Holmes et al. (2004), but Teneggi et al. (2013) also describe a social modulation, showing a link between sensimotor processing and social cognition at neuronal level. Furthermore, Brozzoli et al. (2010) find action specific remapping of the human peripersonal space investigating pointing gestures in comparison to grasping. Scorolli et al. (2014) in addition ascribe humans a sophisticated method to interpret the intentions of social postures and gazes in another's peripersonal space.

---

[1] `http://aldebaran.com/en/humanoid-robot/nao-robot` (visited: Wed 29th Apr, 2015)

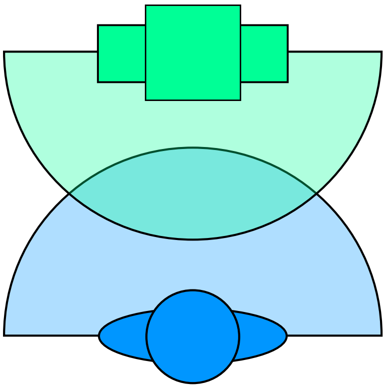


Figure 2.5.: Above view of two interacting individuals. Two dimensional projections of their approximately spherical peripersonal spaces are illustrated as transparent semicircles. The so-called *interaction space* is formed by the overlap of both individual spaces.

Models of the reaching space have also been developed for *social robots*. Chinellato et al. (2011), e.g. learn a mapping between visual input and joint configuration to generate a biologically inspired model of peripersonal space. Nguyen & Wachsmuth (2008) similarly integrate tactile and visual stimuli of a virtual agent to learn such a model by self-exploration. Also, peripersonal space models can for example help to structure an active vision process and thus facilitate object recognition tasks for robots as demonstrated by Goerick et al. (2005).

Nguyen & Wachsmuth (2011) apply the idea of *aligning* representations between humans and machines to the concept of peripersonal space. They relatively accurately estimate the humans peripersonal space by projecting an artificial agents own model to the interlocutor. Similar to the definition of o-spaces by Kendon (1990), they call the overlap of both peripersonal spaces, i.e. the area both agents can act in, *interaction space*. Most recently, Renner et al. (2014) observe a roughly linear degradation of pointing gestures with distance in a shared interaction space that suggests a tendency for humans to respect the other's territoriality in such a scenario. In this thesis, the notion of interaction space accordingly refers to the three dimensional intersection of two agents' peripersonal spaces. Please refer to Figure 2.5 for a sketch of the interaction space as an overlap of two individuals' peripersonal spaces.

2.2.4. Influencing Spatial Configurations in HRI

Hüttenrauch et al. (2006) argue that only simple parametrized models of proxemics and f-formations do not suffice to enable sophisticated social behavior on a robot. According to them, robots have to be equipped with active strategies of altering spatial configurations. Green & Hüttenrauch (2006) therefore introduce *spatial prompting* as “active strategies of the robot that are intended to influence users to position themselves in a way that is advantageous for further communicative actions”. Peters (2011) also describes small steps and body movements of humans as communicational cues intended to reposition a mobile robot. More precisely, Peters (2012, Pg. 172) defines *motionings* as “abrupt motions of one or two steps or sways, bounded by short pauses and are repeated up to five times.”

During the course of this thesis the idea of spatial prompting is extended and transfered to spatial prompting gestures in the *interaction space*. Especially in Chapter 5.5, the term spatial prompting will be utilized in the sense that it refers to a robot’s active *spatial signals*, although the strategies very much correspond to the character of small-scale motionings. In particular, robot movements that solely aim at the repositioning of a humans hands are considered as prompting gestures.

2.3. Summary

In conclusion, this chapter links related work from social sciences to the field of *HRI*. It puts this work into context on the one hand by describing important prior knowledge and how this is used in current social robotics research. On the other hand, it also limits the scope of this work by explaining how certain notions or concepts are interpreted and used throughout this work.

Section 2.1 demonstrates how general theories from social sciences are put into the context of robotics research. With *alignment*, the concept of an automatic adaption process that also facilitates human-robot communication is introduced. Social *signals* are subsequently presented as a basic concept of information transfer. As a last point, the importance of expectations in asymmetrical interactions as with robots is highlighted.

These concepts lay the foundations and guidelines for the development of interaction strategies in this thesis. Social signals sent by a robot are designed to meet a user's expectations so that they support an aligned representation of the situation leading to an improved interaction.

Section 2.2 investigates in greater detail the communicative channels used in this work to improve an interaction. As a prerequisite for raising a robot's social competence by providing it with mechanisms of spatial awareness, the link between interpersonal distance and social meaning is discussed in the form of *proxemics* and *F-Formations*. Also, with the concepts of *peripersonal space* and *interaction space*, representations of space on a smaller scale are reviewed. Finally, *Spatial prompting* as a method to spatially influence another is demonstrated.

With the characteristics of spatial behaviors in humans (and robots) this chapter provides the necessary theoretical background for further investigation of spatial awareness for a *social robot*. The next chapter specifies research questions on top of these findings and explains how they are approached and operationalized in this thesis.

3. An Integrated Concept of Spatial Awareness

This chapter presents an integrated concept of spatial awareness for a social robot. At first, Section 3.1 formulates the overarching concept of how spatial awareness can foster a better understanding of the robot and therefore lead to an enhanced interaction. Secondly, in order to analyze an entire encounter between human and robot, it is decomposed into single phases in Section 3.2. Section 3.3 displays possible strategies for each of these phases that contribute to the overarching idea of spatial awareness. Section 3.4 then explicitly addresses the joint *interaction space* with a newly developed approach to structure it socially. At the end, Section 3.5 concludes the chapter with a brief summary of the proposed concept.

3.1. Proposition of Impact

As suggested in Chapter 1, the robustness of a robotic system can be improved if it is equipped with social features that conform to user expectations. In order to reach this goal, the robot must be able to interpret spatial configurations and signals. It is claimed in this thesis that reasonably coordinated strategies which are based on such knowledge can improve the interaction by giving the user subtle hints that reveal the robots inner state, for example its current intention and focus of attention. Thereby ambiguities and insecurities towards the robot are reduced leading to a better user experience. In other words, the primary hypothesis in this work can be stated as follows:

Hypothesis H *Well-matched spatial communicative strategies can guide a naïve person through an entire encounter with a social robot and thereby foster mutual understanding which leads to an enhanced user experience.*

In order to effectively support mutual understanding, such strategies have to be appropriate to the current situation and not interfere with other behaviors. Additionally, each signal affects subsequent parts of the interaction which requires them to be adjusted to each other. In this work they are therefore explicitly designed in order to help the robot to be perceived as (more) socially intelligent (cf. Fong et al. 2003a; Chap. 1). More specifically, the following criteria serve as a guideline:

- Signals should facilitate mutual understanding by creating an *aligned* state between human and robot (cf. Chap. 2.1.1)
- Signals should be meaningful and efficient (cf. Chap. 2.1.2)
- Signals have to be obvious, co-ordinated, and meet the user's expectations (cf. Chap. 2.1.3)

As a consequence, strategies that are presented in Section 3.3 embody minimal, subtle, and nonverbal *spatial signals* that aim to be human-like and therefore naturally understood by an interaction partner. They try to positively influence the quality of an interaction as well as the perceived properties of the robot during a communication with a human interlocutor.

3.2. Modeling Encounters

In order to accompany the whole situation, a dialog between two agents is considered as embedded in an entire encounter that already begins when both partners are still apart from each other. This thesis therefore fills the gap of far-away interaction to closer ones by persistent investigation and manipulation of spatial constellations.

A typical interaction as depicted in Figure 3.1 is rooted in a situation where two persons are yet unaware of each other. There is a certain distance between them, and they are possibly not in the same room. During the course of this thesis, such a configuration is referred to as the idle stage. The approaching phase is induced when both notice each other and initiation signals are sent. During the approaching phase, one or both partners shorten the distance between them until a comfortable communication distance is reached (cf. Hall et al. 1968).

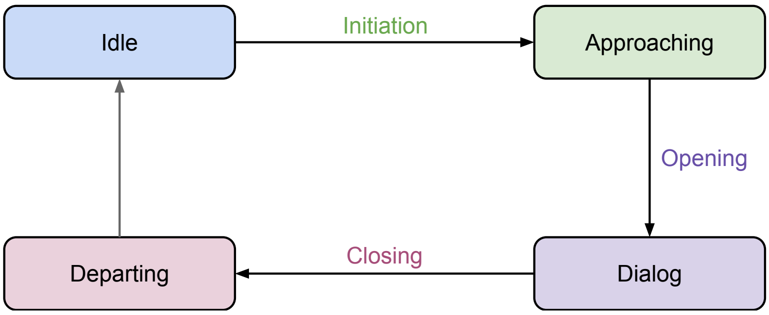


Figure 3.1.: Sequence diagram of two persons engaging and disengaging in a dialog. The different stages (depicted in bubbles) are triggered by specific signals.

At this moment, the conversation is opened with the first words being spoken. Usually, one of the interactants starts with a greeting phrase which is answered by the other one. Such a categorization of mutual engagement roughly resembles the suggestion of Kendon (1990, Pg. 202), who describes greetings of arrivals at a garden party to generally include "a pre-phase of sighting and announcement, a distance salutation, an approach phase and a close salutation".

Afterwards, the actual dialog phase begins. The conversation is eventually closed with farewell words by both partners. Finally, both partners disengage and enter the departing phase where their distance increases as either one of them or both are leaving. After a successful disengagement, both arrive again in the idle phase without further signaling.

3.3. Outlining Spatial Strategies

This section considers a novel holistic way to accompany a whole human-robot encounter with *spatial signals*. For this purpose, a portfolio of strategies that enable a *social robot* to be spatially aware of its interaction partner are presented. The upcoming six claims postulate strategies that enable the robot to better reveal its own state as well as to respond to subtle hints by the user with the help of spatial signals and thus contribute to Hypothesis H. In consequence, the entire interaction benefits from a better mutual understanding between human and robot.

In the next part, specific behavioral strategies for a social robot are introduced, in order to demonstrate their implication as well as to anticipate their influence on the human interlocutor. Please also refer to Chapter 2.2 for detailed theoretical background on anticipated effects described here. In chronological sequence of the human robot encounter (Fig. 3.1), the following aspects of spatial awareness are proposed to guide a human through each phase of the interaction.

3.3.1. Leading into Interaction

This thesis claims that entering a face-to-face dialog between human and robot can be made more convenient by explicitly addressing the key signals of initiation and opening as well as continuously monitoring the approaching phase. In greater detail, the following three interaction strategies support an integrated opening of a dialog from the first contact:

Claim I (Initiation) *Signaling the robot's availability from afar can lead from idling into noticing each other and an approaching behavior.*

According to Kendon (1990, Pg. 165), there is a multitude of ways to initiate an interaction from afar but a common point of origin among them is the identification of an individual as the person to get in touch with. An initiation signal sent by the robot therefore particularly aims to draw attention to it in order to lead into the approaching phase. It reduces a humans uncertainty whether the robot is turned on or off, as well as communicating the ability and willingness of the robot to interact.

Claim II (Approaching) *Gradually attending a human leads into a seamless transition from distant to close-up interaction.*

Both, the description of human greeting by Kendon (1990, Chap. 6) as well as the formal notation of social distances by Hall et al. (1968) allow for the conclusion that approaching each other consists of a gradual process of mutual involvement. So distance dependent incrementation of robot attention towards the human is believed to encourage further approaching behaviors while leaving it open for the human to pass on. The human is assured to be noticed by the robot and is also slowly familiarized with movement patterns of the robot.

Claim III (**Opening**) *Pro-active robot greeting at a socially appropriate distance effectively opens up a dialog.*

Opening a dialog between humans underlies many factors such as interaction history, social status and according to Schegloff (2002) in general is a complex process. Pro-active robot behaviors, as for example presented by Lütkebohle et al. 2009a, have been shown to guide a user through an interaction, so for opening it is assumed to have a similar effect. With regard to proper *proximity* and *orientation*, a self-initiated greeting is believed to expose the robot's verbal capabilities and also acts as an obvious entry point for the dialog phase.

3.3.2. Guiding through Conversation

Human dialog is already accompanied by iconic and lexical gestures (cf. Kendon 1997) that contribute to the robustness of conveyed information. Moreover, complementary motions, as for example proposed by Jung et al. (2013), can provide additional benefits for an interaction between humans and robots. With regard to spatial configuration, the following approach is believed to accompany multi-modal dialog in order to support Hypothesis H:

Claim IV (**Dialog**) *Distinct gestures and spatial prompting can help to resolve ambiguities and conflicts in the interaction space.*

It is claimed that gestures cannot be treated independently from social arrangements in the interaction space. Spatial prompting strategies as introduced by Green & Hüttenrauch (2006) could therefore be applied to the domain of gestures in order to enable a robot to react to physical presence of the human. Chapter 3.4 accordingly investigates the social properties of the interaction space between human and robot. Gained insights result in a model that lays the foundation for spatial prompting and other supportive gestures during a human-robot conversation.

3.3.3. Ending an Encounter

In order to conveniently lead through an interaction, the ending has to be considered as well. Especially in robot systems that are acting slowly, confusion could emerge whether there is more information being provided or the utterance is over (cf. Krämer 2005). Explicitly closing the interaction instead of waiting for the human to act as well as actively leading into a departing phase therefore is considered to have the following beneficial effects:

Claim V (Closing) *Mixed-initiative farewell strategies effectively close a dialog between human and robot.*

Similar to the opening of a dialog, especially in human-robot interactions, the exact moment when an interaction is coming to and end is not always clear. Mixed-initiative strategies, as for example presented by Peltason et al. 2009, have been shown to require less clarification and therefore also qualify as a method to close the dialog appropriately.

Claim VI (Departing) *Signaling robot standby behavior leads to the human disengaging and departing the robot.*

Actively signaling disengagement after the dialog has been closed further clarifies the end of an interaction and leads to mutual disengagement. As the robot does not show attention anymore, such behavior is believed to finalize the current interaction (cf. Ghosh & Kuzuoka 2013).

3.3.4. Implications for a Social Robot

To sum up, the previous section elaborates on Hypothesis H in greater detail. It proposes a social robot that is attentive to spatial configurations and signals sent by the human. In consequence, it makes use of this information to be able to emit non-verbal behaviors itself. Such behaviors then can act as a guidance for the partner in order to consolidate the interaction in terms of robustness. A claim for every step of a whole human-robot encounter that contributes to the hypothesis has been proposed. Each of these claims provides a method to enhance an interaction by providing a strategy for the robot that further reveals the its inherent state, features, or possibilities with the help of *spatial signals*.

At first, a convenient way to lead a human into close encounter with the robot is suggested. Theoretical concepts from social science provide a rich source of information for structuring the social meaning of spatial configurations during the early phases of the interaction. As comparable effects also occur in *HRI*, initial robot strategies can be derived from assumed effects in the human counterpart. Secondly, the concept of *spatial prompting* is anticipated to be applicable in solving coordination problems during conversation. Section 3.4 therefore presents a novel approach for structuring the interaction space. Lastly, ending an encounter is addressed with spatial strategies incorporating active robot behaviors that are intended to function complementary to the opening strategies.

3.4. A Social Structure of Interaction Space[1]

As described in Chapter 2.2, humans coordinate actions in their surroundings with the help of a well structured and social representation of space. Albeit most of the related work concerning this topic regards spatial configurations and signals on a whole-body scale, social effects also occur on as smaller scale (cf. Lloyd 2009). Such effects however still have not been explored extensively in interactions between humans and *social robots*. In face-to-face encounters with social robots, this work proposes the presence of the human in the common *interaction space* to have a social meaning for the interaction. Like in larger scale interactions, it is assumed that humans ascribe the robot some degree of spatial competence and therefore expect it to consider this kind of information in its actions.

Thus, it is important for the robot to represent human presence and activity in the interaction space. It is suggested to structure a robot's representation of space in a similar way to the humans *peripersonal space*. By explicitly annotating this space, a robot is then supposedly able to recognize activity from humans and (re-)act accordingly. A record of interaction histories can help the robot in disambiguation tasks such as which object to grasp next, or what to reference next during dialog.

In order to evaluate in detail how to represent human activity in such a model as well as to infer possible behavioral patterns, this section investigates the expectations of humans with regard to grasping preferences. With regard to the concept of interactive *alignment*, it is assumed that humans project their own spatial representations to the robot as a basis for understanding the robot (Chap. 2.1). To support this theory, three assumptions have been addressed in an on-line video study:

---

[1] Parts of this chapter have been previously published in Holthaus & Wachsmuth (2012)

Handedness:	Humans attribute the robot handedness, i.e. assume the robot to be right-handed like the majority of human population on earth (cf. Perelle & Ehrman 1994).
Distance:	Humans have distance-dependent expectations on the robot's behavior in peripersonal space, so that the robot prefers closer items to those farther away.
Presence:	Humans expect a robot to be influenced in terms of territoriality and focus by the presence of human hands in the interaction space.

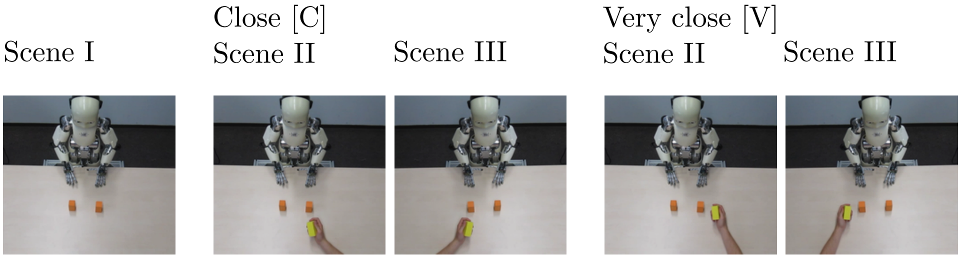
The following sections at first describe the survey setup as well as the different conditions and participants. Secondly, the results of the questionnaire are presented and related to the above hypotheses. Finally, conclusions are drawn for modeling social characteristics in the peripersonal space that enable a humanoid robot to exhibit appropriate interactive behavior.

3.4.1. Probing Spatial Interaction Strategies

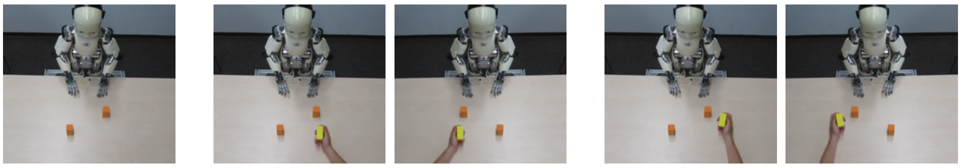
To investigate human expectations towards the social capabilities regarding the interaction space of a humanoid robot, an on-line study (cf. Appx. A) containing images and videos of an interaction situation between iCub (cf. Metta et al. 2010) and a human has been conducted.

In the course of the questionnaire, personal data such as age and occupation is asked at first. Then, the actual scene with two equal objects on a table between a human and the robot is presented from the participants' perspective (Scene I) as a still image. The introductory text states that during the interaction, human and robot should grasp and inspect objects on the table after each other. Participants are provided with the information that currently it is the robot's turn and below the picture they have to choose which object the robot should pick and inspect in their opinion.

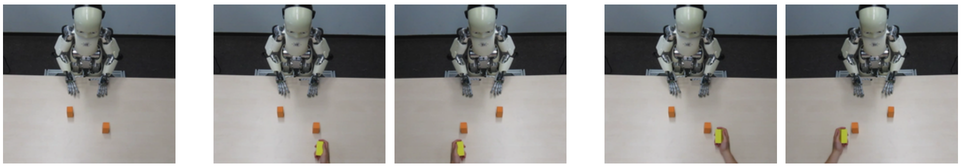
On the next page, a video of the same situation is presented, in which a human hand enters the scene and moves towards one of the objects (Scene II). The hand holds a third object and that is turned around inspected by the human. Again, users have to choose which object the robot should pick in



(a) Snapshots from the middle [M] condition.



(b) Snapshots from the robot [R] condition.



(c) Snapshots from the human [H] condition.

Figure 3.2.: Object and hand positions presented in the video study grouped by object placement. In (a), both objects have the same distance to the robot. In (b) and (c) the object where the manipulation takes place is either closer to the robot or to the human. From left to right, one can see the initial image without a human hand (Scene I), then the human hand present (Scene II) as well as the corresponding mirrored scene (Scene III) in the central columns. The rightmost images show alternate scenes II and III, i.e. in the very close human hand condition in normal and mirrored variant. Some images © 2012 IEEE.

its next move. To avoid effects of handedness, exactly the same situation has been presented a third time only flipped horizontally (Scene III). The first three columns of images in Figure 3.2 give an impression on the order of display. Mirrored scenes have been shown in random order, i.e. in approximately half of the cases scene one and two, i.e. picture and first video, were mirrored and instead of only the second video.

In order to clarify which object users want the robot to pick depending on the distance, the positioning of objects has been varied in the following three conditions:

- Both objects lie in the middle of the table with the same distance to the robot and the same distance to the human observer (Condition [M], Fig. 3.2a).
- The object where the manipulation occurs lies closer to the robot than the other one (Condition [R], Fig. 3.2b).
- The object at which the human hand appears is closer to the human than the other one (Condition [H], Fig. 3.2c).

Depending on a second condition, the manipulation takes place either in the close surroundings between the human and the object (Condition [C], central columns of Figure 3.2), or in very close distance almost besides the right object (Condition [V], rightmost columns of Figure 3.2). A third condition constitutes the ending of the video: Either the hand leaves the scene (Condition [L]) or it stays in the same spot after investigating the object (Condition [S]). Different endings are omitted from the graphics.

On the later pages of the questionnaire, people also answered whether they think that the robot is capable of grasping and recognizing the object, and whether the object's distance or the human gesture actually influence the robot's choice.

A total of 154 people have participated in the questionnaire, of which 14 have been excluded because they did not accept the privacy statement plus another three who did not answer all questions. 51% are females and the participants' age ranges from 23 to 61 with a mean of 30.5 years. The questionnaire was available in German and English, with each person fluent in at least one of the languages. The average self-assessed German knowledge is 3.89 on a five-point scale (0-4), as introduced by Likert (1932).

The participants' occupation varies greatly although many of them are either students (35%) or scientific staff (23%). Most participants have a common technical understanding with an above average computer experience of 2.99 out of 4. The average experience with robots is lower (1.46), so most of them are naive to the actual experiment.

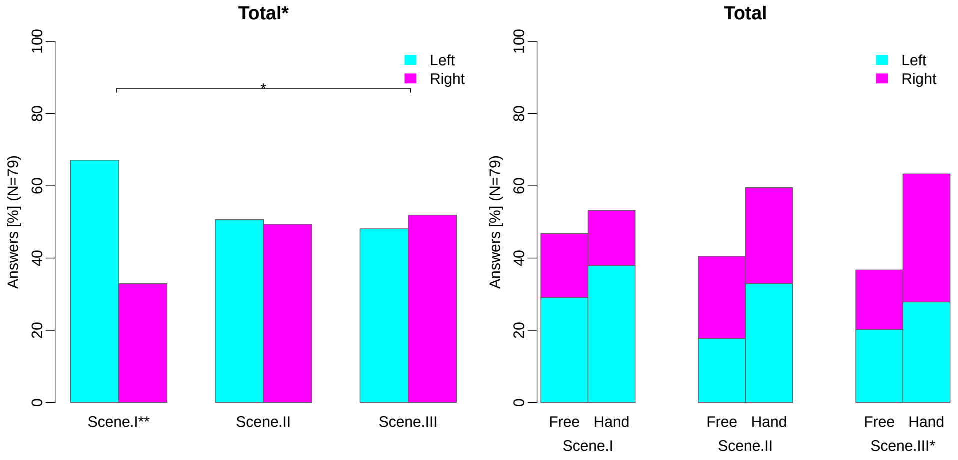
3.4.2. Evaluation of Human Expectations

Answers to every decision question between the two objects as well as on the influence of distance and gesture have been evaluated with the help a χ^2 goodness of fit test (cf. Pearson 1900) against equal probability to distinguish preference for an answer from chance. Also, a χ^2 test of independence (cf. Pearson 1900) has been conducted between the conditions to clarify whether the differences influence the participants' choice.

Almost every participant (96.7%), thinks that the robot is capable of grasping the objects on the table, whereas only 57.6% believe that it can recognize the human gesture. Because the opinion whether one believes in the functionality of the robot or dislikes it heavily influences answers to other questions (cf. Riek et al. 2010; Takayama & Pantofaru 2009), people with a negative vote on either of these questions have been excluded from the rest of the analysis. In total, 79 participants who ascribe the robot perceptual and moving capabilities remain to be evaluated.

Side and object choice Figure 3.3 shows an overview of which object participants want the robot to grasp. In Figure 3.3a, results are grouped by the position of the object summed over all conditions. Participants answers significantly depend on the scene ($\chi^2 = 6.8, df = 2, p < 0.034$). Also, a highly significant preference for the left object can be observed in the first scene ($\chi^2 = 9.2, df = 1, p < 0.003$). Please notice that in this scene (Scene I), there's no hand visible at all, as it is a still image. In the later videos (Scenes II, III), this bias cannot be reproduced anymore, each side is equally chosen. Answers in the third scene significantly differ from those of the fist one ($p < 0.025$).

Part 3.3b describes the results from a different perspective. Answers in this plot are grouped by the place, in which the human hand is active, independent from the object position. The hand could appear either on the right side (default) or on the left side (mirrored setup). Again, in the first

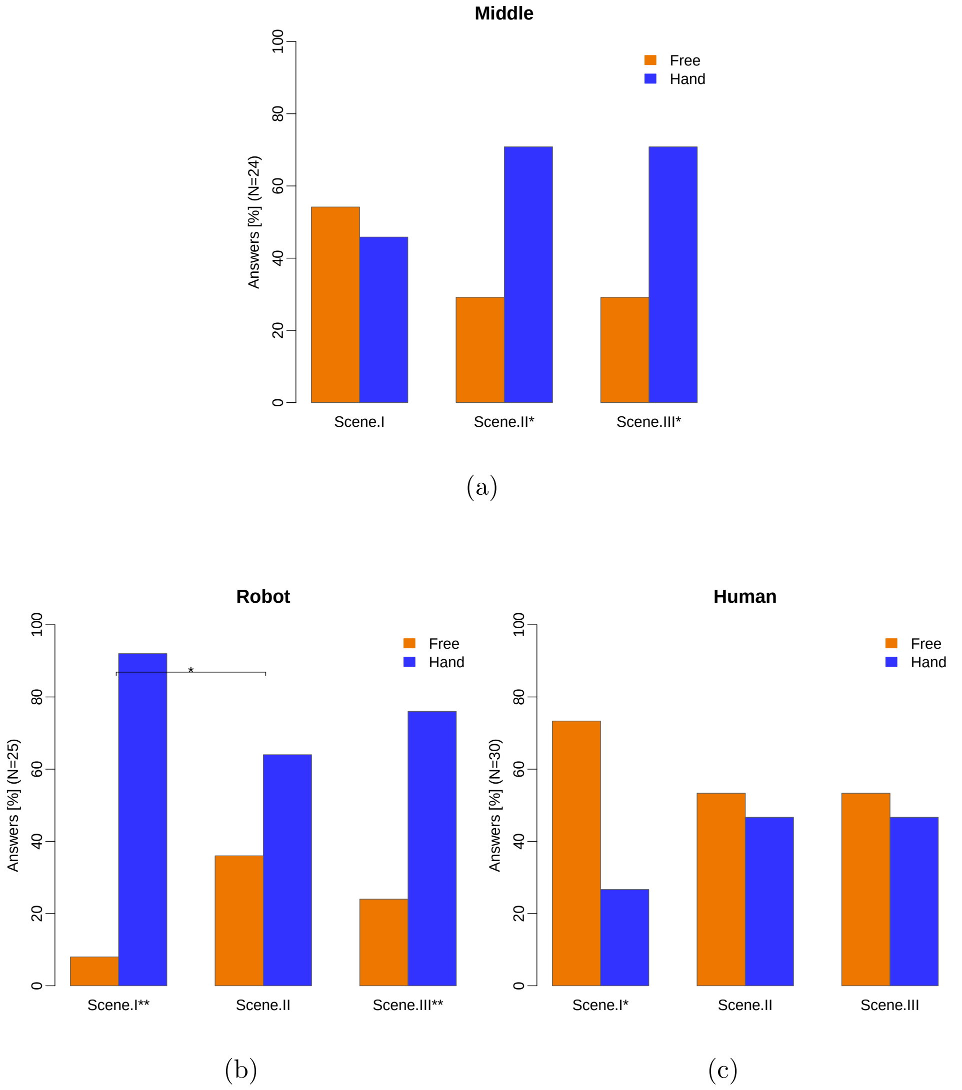


(a) (b)

Figure 3.3.: Overview of the participants' suggestion for the robot in the three scenes. In (a), it is depicted whether the right or the left object has been chosen. (b) contains choices grouped by the object, i.e. whether a hand will be presented nearby the object or not. Significance levels ($* := p < .05$, $** := p < .01$) of the goodness of fit against equal probability are given below the columns. Levels resulting from the independence test are given as bars between the columns and at the title for the complete dataset. © 2012 IEEE.

scene, no hand is visible. Nevertheless, one of the objects lies at the position where the hand appears in the later videos. No significant differences are observable in the first and second scene. A non-significant tendency in the second, and a significant one in favor of the object with a hand in its surroundings is visible in the third scene ($\chi^2 = 5.6, df = 1, p < 0.018$).

Positioning of the object The distance between the object and the robot has significant influence on the object choice of the participants, as depicted in Figure 3.4. For the first scene, a (highly) significant preference for the object that lies closer to the robot can be observed in the robot [R] (Fig. 3.4b, $\chi^2 = 17.6, df = 1, p < 0.001$) and human [H] (Fig. 3.4c, $\chi^2 = 6.5, df = 1, p <$



(a)

(b)

(c)

Figure 3.4.: Participants' object choice in the three scenes with regard to appearance of a human hand grouped by object position. In (a), (b), and (c) answers from the distance conditions (robot, middle, and human) are plotted separately and the difference in object choice is individually compared among the three scenes. Significance levels ($* := p < .05$, $** := p < .01$) are given along the columns, as bars between the columns, and at the title. © 2012 IEEE.

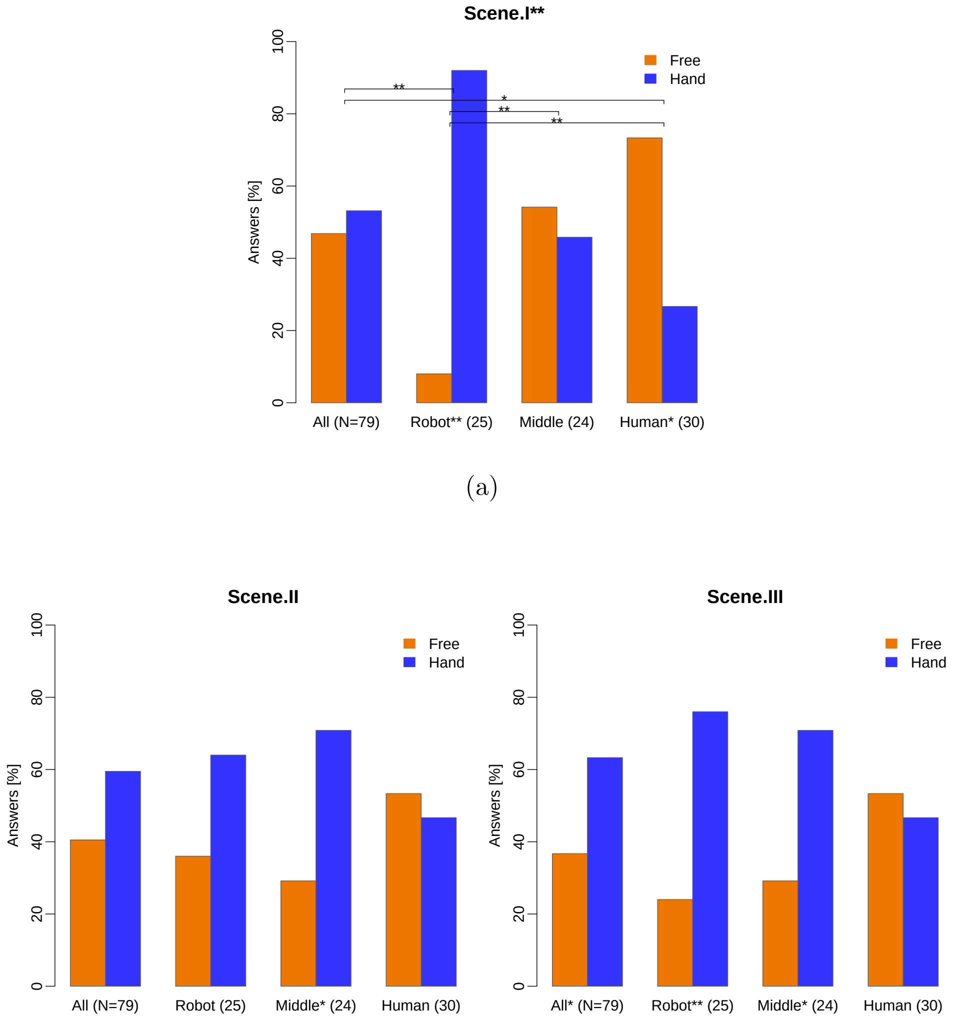
0.011) conditions. Consequently, if both objects are equally far away from the robot ([M] condition, Fig. 3.4a), the answers cannot be distinguished from chance in the still image.

Figure 3.5 describes answers to the same questions as Figure 3.4 but this time grouped in a different way. A comparison of the particular conditions displays the user preference in each scene separately. Figure 3.5a shows that the answer for the still image in Scene I is highly dependent on the condition ($\chi^2 = 24.1, df = 2, p < 0.001$) and that all conditions differ significantly from each other except the middle condition from all results. The presence of a hand influences the choice in Scene II and III as indicated by Figures 3.5b and 3.5c. Conditions cannot be distinguished from each other because all results (except the human condition) show a significant favoring of the object closer to the hand in both cases.

In the human [H] condition, a significant preference in the first scene disappears in the videos with a hand. People choose an object more often if there is a human activity in the surroundings than in the first observation. The opposite holds for the robot [R] condition where people select the free object more often, at least in the first scene with a hand. In the second video, people again prefer the object close to the robot. In the middle [M] condition, a hand nearby the object produces a significant difference from chance in favor of the object to be chosen in scene two and three ($\chi^2 = 4.17, df = 1, p < 0.042$).

Positioning and presence of the hand The results further show that participants in scene three prefer an object significantly more if the hand stays nearby the object than if it leaves the scene (Conditions [S, L], $\chi^2 = 4.6, df = 1, p < 0.033$), although there are also tendencies in the latter case. Trends are visible in scene two as well, but no significances can be observed.

Similar results can be found when comparing the position of the hand in relation to the object (Conditions [C,V]). If the hand is very close [V] to the object, no significances can be found. On the contrary, if the hand is in the close surroundings of but behind the object from the robot's perspective [C], people choose the object significantly more often in scene three ($\chi^2 = 4.8, df = 1, p < 0.028$). A tendency is observable in scene two.

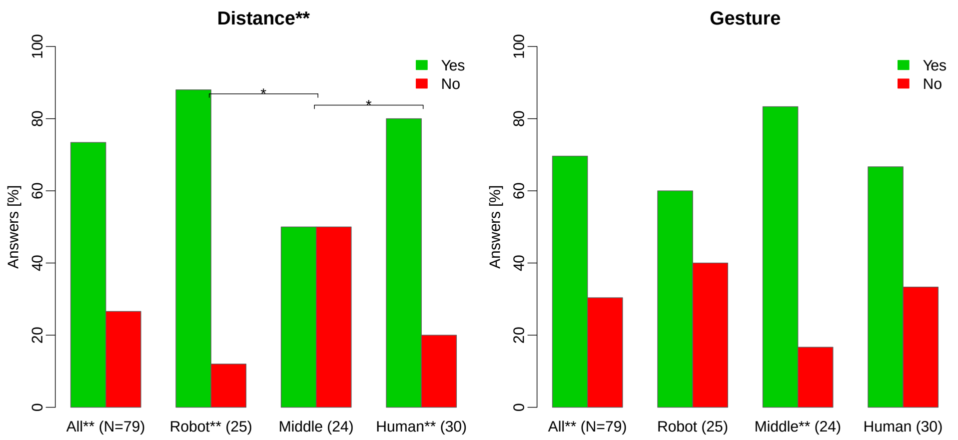


(a)

(b)

(c)

Figure 3.5.: Participants' object choice in the different placements with regard to appearance of a human hand grouped by scene. In (a), (b), and (c) answers from the three scenes are plotted separately and the difference in object choice is individually compared among the placement conditions (robot, middle, and human). Significance levels ($* := p < .05$, $** := p < .01$) are given along the columns, as bars between the columns, and at the title. © 2012 IEEE.



(a) (b)

Figure 3.6.: Participants' opinion about the influence of (a) the object distance and (b) the human gesture on the robot's choice grouped by object placement (robot, middle, human). Significance levels ($* := p < .05$, $** := p < .01$) are given along the columns, as bars between the columns, and at the title. © 2012 IEEE.

Decision questions on influences In the questions after all scenes, people vote significantly for the distance to have an influence on the robot's choice equally among all conditions ($\chi^2 = 17.3, df = 1, p < .001$). One exception from that is the middle condition, where both objects have the same distance from the robot. Answers in this case differ significantly from the other conditions and it could not be distinguished from chance whether participants voted the distance to have an influence or not (cf. Figure 3.6a).

The gesture instead is not significantly believed to have a strong influence on the robot's choice in all conditions, see Figure 3.6b. While the overall influence is approved by the sum of all votes ($\chi^2 = 12.2, df = 1, p < 0.001$) and especially in the middle condition where distance cannot be the deciding factor, answers from people in the human and robot condition showed a trend but did not produce a significant result.

3.4.3. Consequences for Spatial Modeling

The above results suggest that a robot's representation of its surroundings should on the one hand model objects in a distance dependent manner during face-to-face interactions. On the other hand, also the presence and history of hands has to be remembered as participants expect human presence to influence the robot's attentional focus. In general, it could be shown that humans tend to project their own spatial representation onto the robot. Regarding the assumptions stated in the beginning of Section 3.4, the following consequences emerge for modeling a robot's peripersonal space.

Handedness Participants expect the robot to be right handed. When people only see an image with two equally distant objects on a table, they expected the robot to grasp the object on the left, which is the robot's right side. As almost all participants are right-handed, their intuitive response seems to project their own preference for the right hand onto the robot. If other factors such as human hands are involved, this first assumption might still be valid, but apparently superseded by these factors.

Distance The object's position has the greatest impact. If one of the objects lies closer to the robot, people initially prefer this one, because it is easier for the robot to reach than the one farther away. This choice again indicates an assignment of a spatial model which is similar to the human. Such distance dependent behavior can be modeled effectively in a spherical representation of the peripersonal space, because decisions can be based on only one single parameter.

Presence Human activity alters the object choice. Distance alone cannot be used as a deciding factor for the robot in an interaction scenario. Answers from the video study clearly showed that also the presence of the interaction partner has to be considered. A human hand on the one hand can raise the decision frequency if the object is at the outer limits of the peripersonal space. On the other hand, it can inhibit a decision for an object that is clearly in the robot's personal zone. The results indicate that the human can function as an attention getter in the first case but also occupy areas for personal use in the latter case.

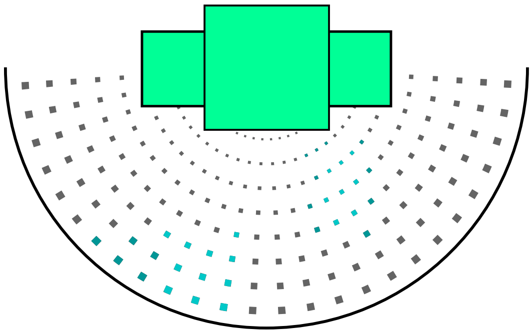


Figure 3.7.: Planar cut at shoulder height through the depiction of the active peripersonal space. A black circle delineates the extent of the individual reaching space and exemplary dots mark the activation grid that stores information about the presence of others. Different shades of teal display the amount of human activity at the respective grid point.

Furthermore, a human does not necessarily have to touch, refer or point to the object directly. In the experiment, the hand holds a third object the whole time which makes a direct reference impossible. Instead, the object is marked indirectly through the positioning of a hand. The decision for the object does also not only depend on the distance between hand and object but additionally on the direction from where the hand approaches the object. If the hand happens to be very close besides the object, the effect is weaker. Possibly, this again indicates a form of occupancy. The ending of the scene also influences peoples choice of the object. Only if the hand stays in the scene, a significant effect is observable. The hand might have been visible for a too short amount of time to be useful as an attention getter. Nonetheless a trend towards the object close to the hand is noticeable in the data.

The above conclusions that are drawn from participants' object preferences are supported by their answers to the decision questions as well. People consciously know and expect the distance as well as the presence of a hand to have an influence on the robot's choice. As stated in Chapter 2.1.3, these expectations should be respected in the robot's behavior, in order to enhance the interaction between naive users and robots in terms of usability.

In summary, all three assumptions have been approved by the survey, as humans attribute the robot handedness, and an awareness of distance as well as territoriality in its own peripersonal space. In order to *align* a robot's behavior to these expectations and increase the robots spatial context awareness, a human-like representation of the robot's reaching space is therefore proposed.

The so-called *active peripersonal space* as depicted in Figure 3.7 could act as the conceptual foundation for more profound reasoning on social activities of a human interaction partner. It is restricted to the reaching space in front of the robot as indicated by a black semicircle. A configurable spherical activation grid stores the amount and points in time with human activity that has been registered by the robot. Due to an angular representation and therefore intrinsically higher resolution of the grid in closer robot distance, it lays a solid basis for distance-dependent social decisions in the robot. Consequently, such a representation is employed as a source of information about human presence in Chapter 5.5 where Claim IV is addressed in a multi-modal dialog that involves robot gestures and therefore bears the potential for territorial conflicts.

3.5. Summary

To sum up, this chapter suggests a comprehensive concept of spatial awareness for a *social robot*. Upon the related research introduced in Chapter 2, it presents a novel way of social intelligence for a humanoid robot that relies on spatial configurations and signals. Section 3.1 for that purpose establishes the overarching aspiration of this work in Hypothesis H. Thereby the research question is given as a proposal how to foster human-robot understanding by supporting the construction of mutual mental models.

In Section 3.2, a holistic perspective of a human-robot encounter, beginning with the first (far away) contact, moving over to interaction opening, and concluding after a joint face-to-face situation, is introduced. For each stage during such a situation, six behavioral strategies that are supposed to support the main hypothesis are presented as claims in Section 3.3.

Special attention is given to the social structure of the *peripersonal space* in Section 3.4. It presents an on-line video study that investigates spatial expectations of humans towards social robots. The survey reveals that humans expect a robot to be right-handed and that it respects distances as well as human presence in the *interaction space*. As a consequence, the concept of an *active* peripersonal space is proposed which is intended to function as a computational basis for spatial interaction strategies at close distances.

In this chapter, already a considerable contribution to the field of social robotics has been provided by investigation the social structure in the active peripersonal space. Furthermore, Hypothesis H and the individual claims for each phase, represent a statement which proposes spatial awareness to be useful for *HRI* by extending communicative capabilities. In order to approve the impact of each singular claim with respect to this statement, presented concepts are in the following incorporated onto an actual interaction scenario which is explained in the next chapter. With the specific scenario as a constraint, each claim is then examined in greater detail in Chapter 5 and finally evaluated in Chapter 6 to draw conclusions on the viability of the hypothesis.

4. The Humanoid Receptionist

The scenario of a robotic receptionist has been chosen in order to address the research question stated in Hypothesis H. It appears to be well suited for the considered phenomena as the way of proceeding covers an entire encounter between an information provider and a questioner. Distant initiation of an interaction as well as the approaching phase can be investigated from the moment someone enters a building or room. Additionally, if the receptionist for example has a ground plan at its proposal, interactions in a shared space can be explored. Besides that, the specifics of human disengagement from the robot are examinable as well.

Besides offering the opportunity to research *spatial signaling*, such a scenario is particularly suitable for an investigation of *HRI*. In the first place, a human in the role of the arriver does not need to have any special training in order to participate in the interaction. Apart from the fact that the receptionist is a robot, the general way of proceeding is mostly intuitive as humans are familiar with similar situations. Today, there are already artificial technical systems providing information almost anywhere, for example ticketing machines. Robotic systems, e.g. outdoor guides (cf. Evers et al. 2014) are emerging, so that they become more common in human daily life. Furthermore, Lohse et al. (2007) name the provision of different kinds of information as a function that is being ascribed to certain human-like robots.

Secondly, due to the conventionality of the a receptionist setting and thus easy to comprehend course of action for naïve users, focus of attention is shifted from the scenario itself onto the robot. While offering the opportunity to research spatial signals, the scenario itself does not alter the meaning of such signals as they have a more general validity and are designed independent from the robot's features as a receptionist. As a consequence, gained insights from a study on the communicative capabilities of a social robot are presumably to some degree transferable to comparable interactive scenarios.

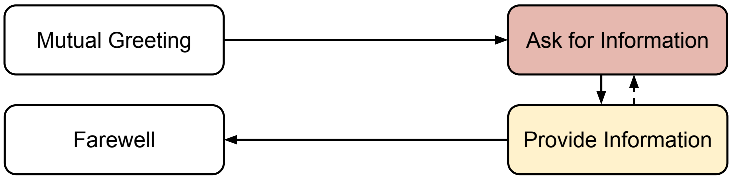


Figure 4.1.: Sequence of a typical interaction between a robot and a human visitor which is shared among common receptionist systems. Human turns are depicted in red, robot actions in yellow. White boxes mark optional steps. In most systems, the dialog phase which is opened by a mutual greeting consists of any number of information requests answered by the robot. Afterwards, the conversation is usually closed by farewell statements.

Furthermore, in a receptionist scenario, a robot can incorporate various digital information sources in order to provide information. It is possible to model a map of the surroundings beforehand as well as to include more advanced knowledge such as office hours from on-line calendars. As another benefit, robot behavior can be modeled straightforwardly due to the relatively limited sequence of actions in such a scenario.

This chapter introduces the basic receptionist that has been realized in the course of this work as an implementation that does not include spatial awareness functions. Instead, it gives insight about the setup and essential functionalities of the receptionist which are needed in order to enable a reasonable interaction. Therefore, it also investigates requirements and constraints for the scenario that emerge from Hypothesis H as well as from the background considered in Chapter 2. With the basic receptionist system as a reference system, it is possible to identify differences in user ratings compared to the advanced receptionist which incorporates spatially aware strategies on top of the here presented functionality (cf. Chap. 5, 6).

At first, other existing humanoid receptionists are described in Section 4.1 on the one hand to approve the feasibility of such a scenario but on the other hand to distinguish their purpose and research question from the ones presented in this thesis. Section 4.2 then investigates requirements for the physical setup. Details on the implementation as well as on the course of action during the encounter is subsequently presented in Section 4.3. Eventually, the chapter is concluded in Section 4.4 with a short summary.

Name	Reference	Research Interest
ASKA	Nisimura et al. (2002)	Speech and Dialog
–	Aleixo et al. (2007)	People Recognition
–	Holzapfel & Waibel (2007)	Dialog Design
	Burghart et al. (2007)	Interaction Patterns
SAYA	Hashimoto et al. (2007)	Head Nods
Valerie	Gockley et al. (2005)	Interaction Design
	Michalowski et al. (2006)	Spatial Engagement
	Lee et al. (2010)	Social Properties
	Salem et al. (2013b)	Politeness and Context
	Makatchev et al. (2013)	Culture and Affiliation
iCub	Holthaus & Wachsmuth (2014)	Spatial Awareness

Table 4.1.: Selection of different receptionist robots: A comparison of functionality and research question.

4.1. Other Receptionists in Social Robotics

The receptionist as an application for robots has been established by a number of research groups investigating different phenomena (cf. Table 4.1). In this section, exemplary robot systems that are implementing the basic logic of a receptionist are introduced. Figure 4.1 displays this common functionality which consists of the human asking for an information, for example the way to an office, followed by an answer of the robot. While all of the presented approaches share the core functionality, each of them lays focus on a unique research question. If the robot is capable of arm movements, the answer is often accompanied by pointing gestures leading the direction. Most setups are also equipped with a dialog that is opened by mutual greetings and ends with farewell statements.

Nisimura et al. (2002) have deployed the humanoid robot ASKA at a university campus, one of the earlier receptionists that is able to produce gestures and is equipped with a verbal interface. They especially address the implementation of a convenient dialog system and investigate the robustness of their speech recognition system for usage in an applied scenario in the real world.

Holzapfel & Waibel (2007) accordingly investigate possibilities to model such a dialog structure for receptionists by reinforcement learning. In contrast, the system presented by Aleixo et al. (2007), albeit also having a speech interface available, it is mainly utilized to improve upon the detection and tracking routines of humans as well as to integrate them into the running system. Both of these approaches to a large degree address the technical aspects of a receptionist system and therefore contribute to the realization of an interaction as depicted in Figure 4.1.

Instead of focusing on the capabilities of the receptionist itself, the work of Burghart et al. (2007) uses the same setup as Holzapfel & Waibel to investigate social phenomena between human and robot. They describe an experimental setting where a social robot guides people to a place they are supposed to deliver parcels to. From such interactions, they infer behavioral patterns of people engaging the robot. They are furthermore able to deduce a transcription method applicable to more general instances of *HRI*. Hashimoto et al. (2007) investigate in a receptionist setup, how head nodding gestures can positively influence aspects of the interaction. They find that as human-likeness, understanding, smoothness and familiarity are increased if the robot expresses nodding at certain predetermined times during the interaction in comparison to randomly timed nods or no head movement at all.

Gockley et al. (2005) introduce Valerie (aka. the roboceptionist), as a permanent setup aiming to research the design of social interaction strategies. They are primarily concerned with long-term effects between the system and untrained users but also are able conduct additional studies with the installation. In this scenario, Michalowski et al. (2006) investigate a model of spatial relationships. Based upon regions around the booth in which the robot is placed, the engagement of visitors is estimated. In their study, they are are experiencing three behavioral patterns of people which they classify as definite interactants, passersby, or still undecided. Lee et al. (2010) additionally confirm that the roboceptionist, although only being equipped with a monitor instead of a head, is being treated as social entity by people.

Salem et al. (2013b) utilize the same receptionist to research effects of robot politeness and context on the perception of the robot. While they cannot find any differences comparing two politeness strategies, the change from a goal-directed to a free-form interaction results in an increased perceived warmth of the robot. More recently, cross-cultural effects in human-robot interactions are being researched by Makatchev et al. (2013), as the robot is capable to express ethnically salient behaviors in a bi-lingual fashion. They are able to show affiliation effects with the robot but cannot finally prove an ethnic homophily.

The setup which is pointed out in the next two sections shares the ideas behind the latter setups. It has been specifically developed in order to investigate the research social phenomena between a robot and a human interlocutor. More precisely, it does not insist to represent the best possible way to implement a receptionist robot. Instead, it consists of an interactive robot system that serves as a starting point to research the effects of an integrated concept of spatial awareness for a *social robot*. Firstly, the physical capabilities of the scenario are outlined in Section 4.2. Afterwards, the basic interactive functionality is displayed in Section 4.3. Strategies that contribute to the research questions are introduced in Chapter 5 immediately after the scenario has been framed.

4.2. Capabilities of a Receptionist Robot

In this section, the technical requirements as well as the physical realization of the basic receptionist is being described. At first, Section 4.2.1 presents the reasoning behind the selection of robot hardware for the scenario. It gives insight which factors are needed as a prerequisite and why the specific arrangement has been chosen. Then, in Section 4.2.2 the concrete realization and final arrangement is being explained.

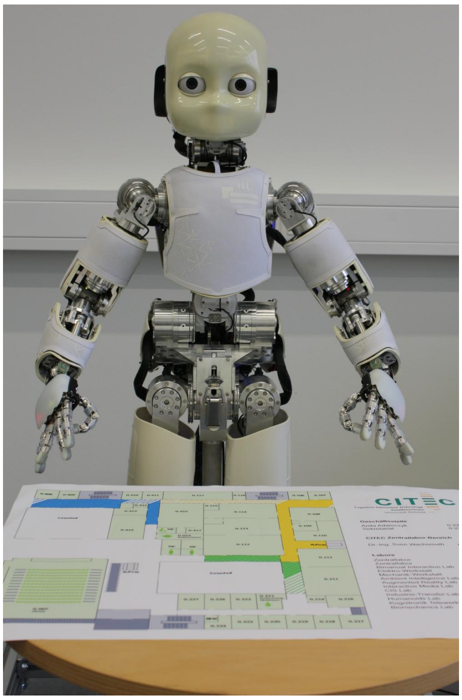
4.2.1. Requirements and Constraints

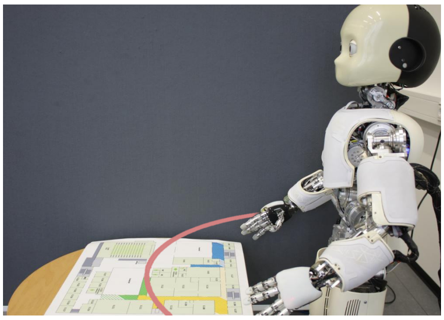
In order to setup a convenient environment for the investigation of social interactions with *spatial signals* between a robot and a naïve user, several requirements have to be considered. In terms of robot appearance and functionality, it is a prerequisite to provide a humanoid look that inherits and exhibits the purpose of social interaction capabilities (cf. Chapter 2.1). Furthermore, the body has to be of an appropriate height, in order to draw meaningful conclusions when spatially relating the robot to a human counterpart (cf. Chapter 2.2). As a final technical requirement, the body has to be equipped with certain movable parts. At first, upper limbs (i.e. arms, hands, fingers) are a necessity for displaying pointing gestures as well as spatial prompting. Besides that, it needs a flexible hip, head, as well as eyes in order to engage in social configurations and display visual attention. With no further requirements on special robot hardware, it is possible to apply the scenario and the spatial awareness concept to a variety humanoid robots with similar capabilities in a convenient way.

As a self-imposed constraint, in contrast to a mobile robot that is moving around freely, a stationary setup has been chosen. Such a decision comprises that the robot is unable to reposition itself because it is fixed at a certain location, namely behind a receptionist desk. While the spatial communicative capabilities of the robot are limited in terms of positioning, at the same time effects that are out of focus for this thesis can be excluded in advance. To be exact, this thesis does neither claim nor attend to address the (certainly important) question of how to communicate navigation goals to a human. With an immobile robot, such processes are immediately ruled out by the scenario composition.

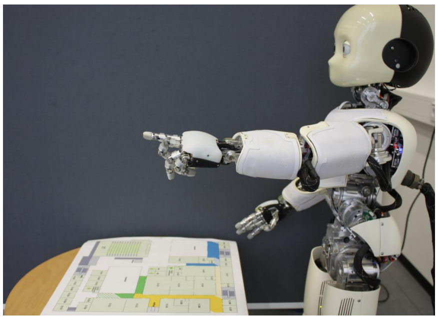
At the same time, a stationary receptionist setup provides the possibility to elaborate on all the claims given in Chapter 3.3. During the approaching of a human, Claim I, II, and III can be examined. Given that *proxemics* and *F-Formations* play an important role during these phases, the scenario is well suited for investigating how to properly use communicative signals such as display of visual attention in order to positively influence the interaction. With the integration of a floor plan into the scenario a shared *interaction space* is created, so that Claim IV can as well be subject to research. For example, if human body parts hinder the robot in pointing towards buildings on the map, and *spatial prompting* strategies might be considered as a socially proper alternative in resolving such a conflict. Similar to opening the interaction, the influence of *spatial signals* can be investigated in addressing Claim V and VI. Explicit disengagement could be considered as an adequate method to lead a human out of the interaction and towards the place she is searching for.

To sum up, a receptionist scenario with a properly designed stationary robot offers the opportunity to effectively research the space in-between human and robot. In such a setup, not the robot's positioning but the interpretation and production of pose and gestures can be investigated if it has the appropriate expressive capabilities at disposal.





(b) Side view: Standby pose



(a) Front view: Standby pose

(c) Side view: Pointing

Figure 4.2.: Setup of the iCub in the receptionist scenario. The robot is fixed on a pole behind a desk with a floor plan on it. (a), displays a front view of the setup, in which facial features of the robot are visible. In (b) and (c) side views of the same scene are given. (b) displays the standby pose as well as an augmented circle, delineating the border of the robot's reaching space. (c) gives an exemplary pointing gesture towards another room.

4.2.2. Installation and Physical Setup

The iCub robot, unveiled by Metta et al. (2010), is incorporated into the scenario in the role of the receptionist (cf. Fig. 4.2). Power supply and Ethernet are wired at the robot's back, so that all computing except motor control can be done on regular PCs that do not have to stand close by. An external 3D camera (not shown) for perception is placed slightly behind the robot above its shoulder so that it can observe attending persons as well as the desk with the floor plan on top. A pair of standard PC speakers resides hidden at the robot's feet. Voice recording is done via a wearable microphone in order to increase recognition accuracy. Optionally, e.g. for demonstration purposes, a directional microphone can be used, which restricts verbal conversation to certain areas directly in front of the robot.

As described by Tsagarakis et al. (2007), the iCub has been developed especially for cognitive interaction research and therefore a lot of effort has been put into a sociable interface as recognizable in Figure 4.2a. The robot's body resembles a full human shape, having a total of 53 degrees of freedom available. With a height of 948mm and 365mm reaching length including arm, hand, and fingers, it is smaller than most adults. When mounted on a pole, its height grows up to 1340mm as it floats above the ground. Because the body is able to bend forwards, despite being fixed on the pole, its reaching limit is extended to approximately 480mm at just above table and thus floor plan height.

The iCub is placed behind a small receptionist desk with a height of 730mm above the ground. A floor plan which measures 380mm x 380mm and is placed 240mm away from the robot on top of the desk (cf. Appx. D), which is almost completely covered by the plan. Figure 4.2b relates the robot's reaching limits to the map by displaying an augmented semi-circle on top of the map surrounding accessible areas. The robot is depicted in the standby pose, ready to welcome visitors. On the contrary, Figure 4.2c shows an exemplary pointing gesture indicating towards another room in the building.

Despite the relatively small size of the robot compared to a normal adult, Walters et al. (2009) report that size differences of about 20cm do not significantly influence approaching preferences of humans towards robots. Human approaching behavior in terms of *proxemics* can therefore be studied deliberately with the iCub, taking into account that results might differ slightly from interactions at exactly the same height. With regard to *interaction space*, the physical capabilities of the robot intentionally divide the map into two approximately evenly sized areas (cf. Fig. 4.2b). There is shared space, where both have direct access as well as areas outside of the robot's scope, leaving the opportunity to investigate different behaviors in those zones.

4.3. Interactive Functionality

The receptionist system is not only realized as a monolithic software. Instead, it connects a number of independent modules, of which many are implemented as standalone processes, e.g. existing software for hardware communication or speech recognition. As an overview, Figure 4.3 depicts how the modules are arranged to create the basic receptionist. Section 4.3.1 at first describes how external modules (orange in Fig. 4.3) are connected to the main component in order to enable perception and action in the receptionist. Afterwards, Section 4.3.2 gives details on the implementation of the indicated functionality inside the executional layer (on a gray background in Fig. 4.3) that relates basic information from the sensors to semantic representations. On this basis the dialog system decides on which action to take and which outputs to produce. Finally, the procedure of engaging and interacting with the basic receptionist is explained using an exemplary course of action in Section 4.3.3.

4.3.1. Communicative Channels

In order to implement interactive functionality, the system has to access input channels like cameras as well as output channels such as robot actuators. Thereby, perception and actuation modules have to be realized separately, as a lot of different hardware is being employed. Thus, inter-process communication in the receptionist is needed, which is enabled on basis of the XTT protocol by Lütkebohle et al. (2011) supported by a memory-like communication framework (cf. Wrede et al. 2004a,b). This allows for an accurate representation of a tasks life-cycle, i.e. monitoring the beginning, duration and ending of behavioral functions. As a consequence, utterances, gestures, and gazes can be executed in a controlled and synchronized fashion. Furthermore, it is possible to monitor and react to errors that occur during the execution of a task.

Visual perception is implemented using the iceWing software developed by Lömker et al. (2006), which captures video data from an external 3d-camera. With the help of a boosting cascade (cf. Viola & Jones 2001), the face of an interlocutor is found in the scene. Hands on the other hand are detected and classified as such using a modified articulated scene model as described by Swadzba et al. (2010), so that it is also applicable in smaller

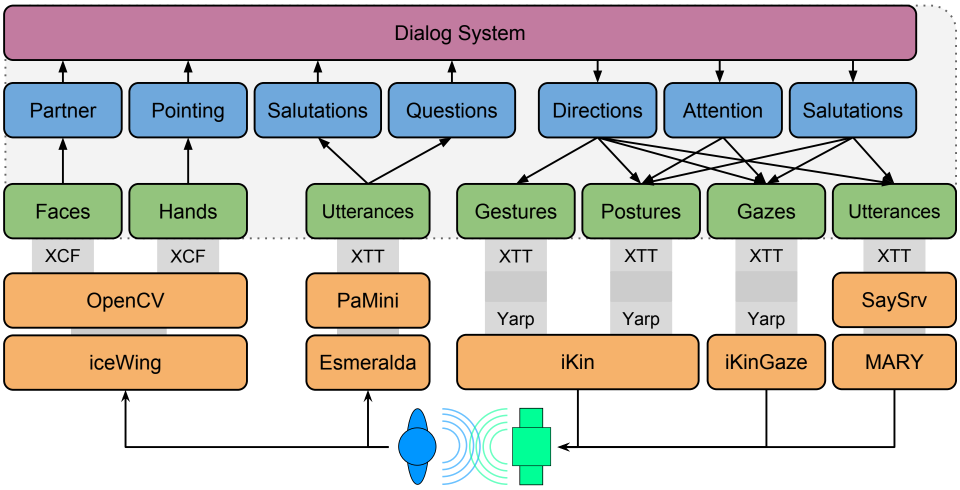


Figure 4.3.: Abstract depiction of the basic receptionist system. In the first three rows (on a gray background), internal representations of concepts important for the interaction are listed. The fixed-decision dialog system (purple) acts as a connection between input and output. Semantic knowledge in the next row is painted in blue. Green bubbles represent the respective units of basic information. Orange bubbles connect each of these units to an input or output channel for communication. Please note how the coloring corresponds to the layers of a technical *alignment* system presented in Figure 2.2 of Chapter 2.1.1. Arrows mark the overall flow of information in the system.

scale conditions. Given a proper calibration of camera systems prior to an interaction, 3-dimensional information about the location of faces as well as human hands are then related to real-world locations in robot coordinates using functions of the OpenCV library (cf. Bradski 2000).

Speech can be perceived via a directed microphone or alternatively a headset that participants have to wear. Also keyboard-based input is possible for testing purposes. Recognition is achieved with the HMM-based Esmeralda software introduced by Fink (1999), which is supported by a fixed grammar (cf. Appx. B). Recognized sentences are used as an input for PaMini, a dialog module which is capable of mixed-initiative (cf. Peltason & Wrede

2010), i.e. triggering of routines by both human and robot. On this basis, system tasks are being executed, that are for example connected to the internal routines for explanation or salutation or helping behaviors.

Movements as an output are solely executed on the iCub robot. For that purpose, it offers several software interfaces for actuating the hardware, such as an inverse kinematics (iKin) module developed by Pattacini et al. (2010) which is utilized in the receptionist to realize pointing gestures. Similarly, gazes can be controlled in a human-like way using coordinated head and eye movements as described by Pattacini (2011). Both modules are being accessed via the YARP middleware introduced by Fitzpatrick et al. (2008). In order to maintain the benefits of the XTT-based task description, an adapter module acts as intermediary between motor actuation and the rest of the software.

As a secondary output channel, speech is being incorporated. It is realized with the text-to-speech synthesizer Mary by Schröder & Trouvain (2003). Utterances are produced independently from the robot and made audible using standard desktop loudspeakers which are placed close to the receptionist desk. A comprehensive list of robot utterance is given in Appendix C.

4.3.2. Internal Representation

The receptionist's main component, which is shaded in gray in Figure 4.3 acts as the executional layer. For that purpose, it is connected via interprocess communication to the input and output module described in the last section. Its functionality consists of the continuous interpretation of perceptual primitives as semantic information and the inference of possible needs for action that finally results in actuator commands.

Due to the need for highly parallel processing, that results from the incorporation of various input and output modules, the software is composed of independent graphs running at the same time. The actual implementation is based on a filter and transform approach as described by Lütkebohle et al. (2009b) and event-triggered by the input sources.

There are two graphs occupied with visual inputs. One transforms the percept of faces into partner locations and continuously updates the internal knowledge about that position. From that information, targets for attention towards the person are inferred that are decomposed into postures and gazes. Similarly, the second graph refreshes the possible pointing locations by taking the information of human hands into account. As a result, the system has the possibility to integrate such knowledge into the dialog at any point in time during the interaction.

Another three graphs realize the verbal interaction and thereby control the flow of actions in the system. One of them addresses the perception and production of salutations, which activates all of the interactive functionality (including attentive behaviors) after mutual greeting and shuts it down after farewell. The second reacts to reset statements and is capable of restoring the starting situation in the robot. The third graph is concerned with the actual functionality of the robot in terms of providing information. It recognizes location questions and produces robot behavior. From the knowledge about the pointing or verbal description of the location, it generates directions for postures, gestures, and gazes. Furthermore, it calculates the correct utterances. As a final step, movement primitives and verbal statements are redirected towards the output modules.

Furthermore, there are some additional graphs that are not directly involved in the interaction. One is realized for displaying debug information on screen and another one for is used for the calibration of the floor plan location to the camera and robot coordinate systems prior to the user engagement.

4.3.3. Course of Action

In the scenario, the receptionist is in idle state and waiting for potential interlocutors. Visitors can approach the robot and ask it in which direction to find certain places. As described in the last section, speech as well as deictic gestures to locations on the map between them can be used as a reference for the inquiry. The robot in all cases uses both speech and gesture in its answer to indicate the correct direction to the visitor. For referencing the target, iCub is capable of referring to the map and locations in the real world.

An interaction with the receptionist robot presented in this thesis follows the same principal procedure as the systems presented in Section 4.1. After a mutual greeting, the main dialog begins. The conversation in essence consists of the human asking for a way and the robot answering by providing the wanted information. Afterwards, the dialog is closed by mutual farewell statements (cf. Fig 4.1). In resemblance to this categorization, Figure 4.4

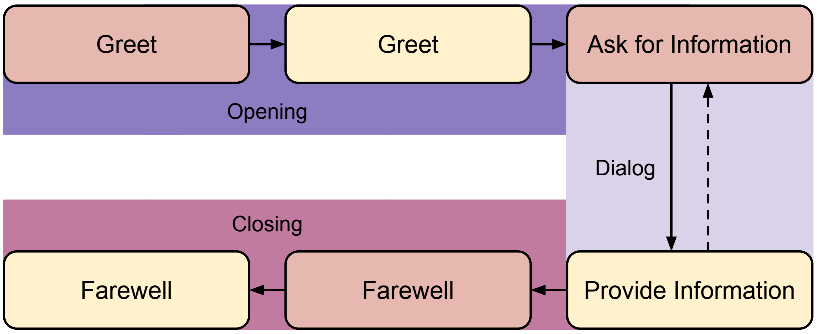


Figure 4.4.: Sequence of an interaction between a receptionist robot and a human visitor in the basic setup. Human utterances are depicted in red, robot utterances in yellow. The dialog phase which is opened by a mutual greeting consists of any number of information requests answered by the robot. Afterwards, the conversation is closed by farewell statements.

relates the dialog turns to the interaction phases presented in Chapter 3.2. Furthermore, it displays the order in which greeting and farewell statements occur in the basic receptionist. Please also refer to Appx. B for a detailed grammar of what the receptionist is able to recognize. Appx. C gives a comprehensive list of utterances that the robot produces, where functionality of the advanced receptionist that is discussed later is marked as (adv).

During the idle and approaching phases, the receptionist resides in a standby pose, slightly leaning backwards and awaiting possible interaction partners. Its gaze is turning in fixed intervals of approximately seven seconds towards randomly chosen targets a few meters away and 500mm around the current focus. By restricting gazes in that way, it can be ensured that no overly confusing or artificial turns occur, e.g. straight towards floor or ceiling. The actual dialog is opened by a human greeting (e.g. “Hello”, “Hello iCub”) followed by the robot answering with the sentences “Hello, my name is iCub. How can I help?”. Simultaneously, the robot leans forward and looks towards the interaction partner, displaying visual attention using its torso, head and eyes.

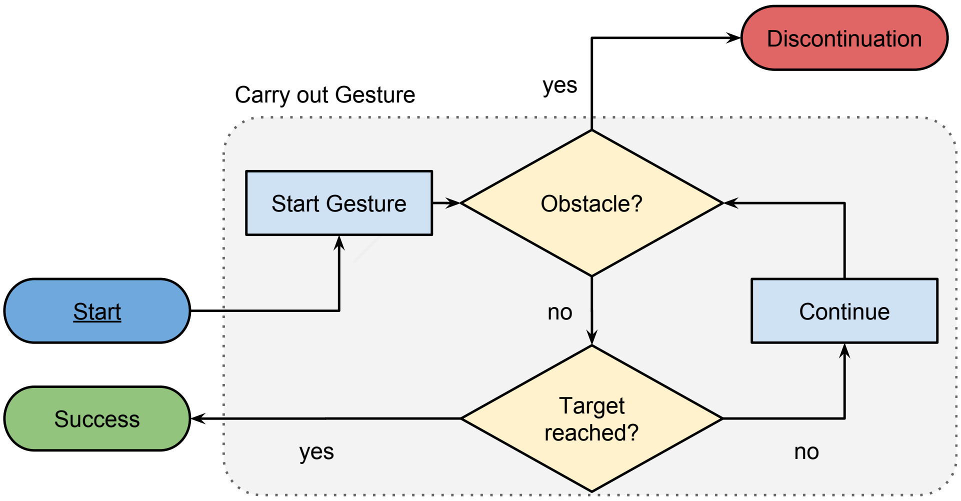


Figure 4.5.: Flowchart of the pointing gesture in the basic receptionist. As soon as a pointing gesture starts, it is checked continuously for obstacles. Immediately upon encountering one, the gesture is discontinued. Otherwise, it is carried on until the target has successfully been reached.

While interacting at the table, visual attention towards the human as well as the randomly occurring gazes are kept consistently. The robot is awaiting questions about locations by the visitor, which it can answer with the help of a pre-modeled map of the surroundings (cf. Appx D). Due to the annotation of the robot on the map, it can relate itself towards other places in real world coordinates. Therefore, not only gestures towards locations on the floor plan can be exhibited but also gestures pointing towards their actual position.

There are two ways a human can retrieve directional information from the receptionist. The first method consists of asking a question after a place, phrased like "Where/What is the [place]?". The second possibility would be to point towards a part of the map and then ask "Where/What is this?". Due to the 3d-vision, which is able to detect the physical map on the desk, the robot can relate a pointing gesture to one of the pre-modeled locations, if it is directed towards one of them.

Answers to questions from the visitor are two-part and involve verbal utterances, gestures, and shifts of attention including gaze and body posture. The first part explains the location in question on the floor plan. For that reason, attention is shifted from the person towards the location by fixating the robot's hip in the direction of the map. In order to prime the user towards the correct location, a gaze in the direction of the location exhibited in the very beginning of the multi-modal gesture (cf. Renner et al. 2014). This gaze is kept during an immediately following pointing gesture towards the same location in order to disambiguate the referential domain (cf. Brown-Schmidt & Tanenhaus 2008, Pg. 676) as well as to increase the accuracy of the pointing gesture as described by Williams et al. (2013). Simultaneously, the receptionist utters "This is the [place].". Thereby, it still points and gazes towards the place's location on the map. During pointing, information about the user's hands is regarded in order to avoid collisions with the human. Figure 4.5 describes in detail how a gesture is carried out and under which circumstances it is possibly aborted.

As a second part of the explanation, the robot says "You can find it in this direction." and at the same time gazes, turns, and points towards the location in real world coordinates using the hand which is closest to the target. After the explanation is finished, the robot again returns to its home position and shifts attention back towards the human. After asking "Can I tell you something else?", it is eventually able to answer additional questions.

Besides explaining the way, the robot also has the capability to offer help on its functionality. Either, the human specifically asks for it by saying "Help" or "What can you do?", or the robot detects a period of more than ten seconds of silence. In both cases, the robot explains its functionality by saying "I am capable of telling you where you can find certain rooms in this building.".

The interaction is closed on behalf of the human. Any of the utterances "Bye", "Good bye", and "Good bye, iCub" terminate the interaction immediately. The robot replies with "Good Bye", or "See you soon." and leaning backwards until reaching the idle pose. Visual attention towards the human is being discontinued, only random gazes are still enabled.

4.4. Summary

In this chapter, the receptionist scenario has been introduced. At first, the feasibility of such a setup has been discussed. Section 4.1 introduces other research that has been conducted in similar scenarios and relates their setup as well as research questions to the ones discussed in this thesis. The amount of similar setups as well as the profound research being conducted with them supports the idea of researching spatial relations in the context of a receptionist. Furthermore, all comparable setups approach a very distinct research questions from the one raised in this thesis.

In Section 4.2, the basic functionality as well as hardware setup is being explained. After investigating the reasoning behind the specific setup of a stationary social robot, the layout of the scene and the physical capabilities of the robot are depicted.

Section 4.3 consists of explaining the interactive features of the scenario. It credits incorporated softwares and algorithms from other developers and explains implementation details. Lastly, the course of action in the receptionist encounter is explained and related to the relevant parts of an holistic encounter given in Chapter 3.2.

In summary, this chapter argues why the scenario is well suited to answer the research question. It is well explored, tailored to the topic and is fixed to some degree but sufficiently adaptable to allow for investigation. The chapter also gives the basic functionality that is needed as a comparison condition in order to explore advanced spatial strategies, that are introduced in the next chapter. Afterwards, Chapter 6 compares both in order to evaluate the benefits of human-like spatial awareness in a social robot.

5. Behaviors for Approaching Human-Like Spatial Awareness

This chapter describes possible advanced interaction strategies for a receptionist robot aimed at enhancing the communication with a human partner. New concepts for spatial awareness on top of the basic receptionist system introduced in Chapter 4 are presented. In the resulting advanced system, every phase of a dyadic encounter between a visitor and the robot is addressed in an integrated fashion based on the propositions in Chapter 3. In the following, they are discussed in order of occurrence during the complete situation.

The first three sections present more abstract concepts prior to the dialog phase. They outline possible strategies regarding the claims introduced in Chapter 3.3.1 with the receptionist scenario as a constraint. Afterwards, the concept as a whole is applied to the overall system and course of action as described in Chapter 4.3. In greater detail, Section 5.1 suggests a straightforward method to signal robot availability from afar and thereby provides an initiation signal for the approaching and later dialog (cf. Claim I). Special focus is then laid on the approaching phase (Claim II) in Section 5.2 which deeply investigates the first steps into the dialog as it evaluates a prototype study on how to signal availability from afar, maintain contact during approaching and eventually enter the conversation. A mixed-initiative approach as the opening signal for the dialog phase (Claim III) is then addressed in Section 5.3. The overall concept for leading into the dialog covering the first three claims is eventually summarized and related to the basic receptionist system in Section 5.4.

The next two sections each combine the proposition of spatial strategies with the application to the actual receptionist system. For that purpose, Section 5.5 investigates how the social structure of an *interaction space* (cf. Chap. 3.4) can be incorporated in order to infer appropriate robot behaviors for the dialog phase. The dialog between the human and the receptionist is thereby enhanced with regard to territorial uncertainty in order to address the issues raised in Chapter 3.3.2 as Claim IV. It thus proposes methods to arrange a robot's arms and body in order to structure the shared interaction space. Additionally, it is described at which points these concepts are incorporated into the existing dialog and gesture system.

Afterwards, in Section 5.6 an approach for effectively closing the dialog by again embracing on the mixed-initiative concepts of interaction opening is introduced. Furthermore, a concept for accompanying the departing phase that relies on actively signaling a lack of interest in conversation as well as a step-wise reduction of involvement is presented. This section therefore provides means to address Claims V and VI and how to apply them onto the basic receptionist.

Finally, Section 5.7 gives a brief summary of the suggested strategies and their expected impact on the receptionist scenario and Hypothesis H. It furthermore gives a precise overview how the strategies are embedded in the software architecture and the course of action of the basic receptionist system.

5.1. Initial Sighting

In order to initiate contact between a robot and a human and lead into an approaching phase of the human towards the robot, this section presents the possibility for an active initiation signal for a receptionist robot. Such a strategy aims at fulfilling Claim I and is employed prior to the beginning of an interaction. How exactly this strategy can be incorporated onto the actual system is described in conjunction with the integrated opening concept in Chapter 5.4.

One possible method of interaction opening consists of a distant salutation, which then leads into an approaching behavior and is for example discussed in Kendon (1990, Chap. 6). Distant salutation is primarily employed between somewhat familiar people and possibly incorporates hand waving or speech from afar. Contrarily to such an approach, it is claimed that in human-robot encounters and especially in a receptionist scenario, a step-wise opening leads to a better user experience. Non-verbal signals are instead employed as means to already reveal functionalities of the robot (cf. Patterson 1982) before the focused interaction itself begins.

As a prerequisite for an approaching phase to begin, Mondada (2009) names the identification of a possible interaction partner as one of the first steps. This identification then leads into an organization of convergence, a constitution of common space and eventually into a new spatial configuration at a closer distance. Because the robot's location is fixed behind the receptionist desk, its greatest opportunity for an initiation signal therefore lies in signaling availability at the very first moment of the encounter. An organization of convergence is thereby triggered by the robot but the execution is left to the human. As a result, the human then possesses the possibility and responsibility to turn and approach towards the robot if an interaction is desired.

Compared to a human partner, the importance of such an initiation signal is even higher with a robot as a receptionist due to the wide range of expectations in conjunction with the inherent uncertainty humans have on the robot and the interaction (cf. Chap. 2.1). A robotic initiative eliminates the burden of a first step for the naïve user and at the same time signals an availability as an interaction partner.

According to Kendon (1990) as well as Mondada (2009), a key signal prior to approaching is the establishment of mutual eye-contact. Subsequently, a short precise gaze towards the human is being proposed as the initiation signal. According to Kampe et al. (2001), it represents a subtle but effective way to signal availability while at the same time being appealing to the human counterpart. Besides clearly signaling that the robot is switched on, a short gaze also reveals the basic attention mechanism of the robot, namely head and eye movements, as well as the capability to recognize a human interlocutor. In consequence, it proposes the robot as a qualified interaction partner but does not force the user into a reaction as opposed to e.g. a more aggressive hand waving.

5.2. Incremental Instigation of an Interaction[1]

This section explores spatial strategies that are applied to the receptionist prior to the dialog phase. The driving idea consists of leading a human smoothly into the interaction after the first contact has been made. Considering that, a first prototype implementing a coherent concept that displays

---

[1] Parts of this chapter have been previously published in Holthaus et al. (2011)

availability as well as a gradually increasing amount of attention is presented in Section 5.2.1. The specifics of this concept are subsequently evaluated with the help of a video pre-study in Section 5.2.2. With the insight of this evaluation as a background, Section 5.2.3 presents slightly adapted strategies for an incremental interaction opening in consequence. The instigation section is eventually wrapped up with the proposal of a spatial awareness concept in Section 5.2.4. After introducing dialog opening in Section 5.3, details on the consolidated implementation in the advanced receptionist as an integrated entry into the dialog phase are given in Chapter 5.4.

5.2.1. Outline of a Viable Approaching Response

This section focuses on the transition between distant and close communication for an interaction opening given that the human has already recognized the robot as a possible interaction partner. For that purpose, it introduces a first outline of a proximity-based attention system for interaction opening, which has been realized as one of the spadeworks for this thesis. Spatial models of *proxemics* and *F-Formations* are applied to a humanoid robot and connected to behavioral patterns displaying visual attention towards the approaching human. The resultant proximity-based attention system which serves as one of the core mechanisms in the holistic robot encounter is introduced in the following.

In order to explore the range of viable behaviors for attending the approaching human, an autonomous system has been implemented. It has been realized in an earlier version of the receptionist which is constituted of an alternative robotic setup. For the torso, it consists of the humanoid robot platform BARTHOC as described by Hackel et al. (2005). Its head has been replaced by the more recent development Flobi (cf. Lütkebohle et al. 2010), which has been explicitly designed by Hegel et al. (2010) to produce social behaviors and human-like feedback as well as to integrate sensor functionality.

Of the 45 degrees of freedom (DOF), only the hip, head, and eyes are being used in this setting (6 DOF). For a visual input, the head is equipped with two fire-wire cameras in the eyes. Since the cameras are attached to the eye-balls, their image always reflects the current view direction of the robot. In the following, the recognition of person locations as well as their respective associations to the attention system are explained in detail.

In contrast to the receptionist that has been introduced in Chapter 4, a much simpler approach has been implemented. Due to the conduction of a video-based study which is restricted to the engaging stages, many features of the receptionist can be omitted. As a major consequence, there is no need to explicitly model the dialog between human and robot. Instead, all functionality is based solely on a non-verbal sensor-actor loop that detects the face of a human using the two-dimensional camera in the left eye of the robotic head. An interaction partner for the robot is detected with a standard face detection algorithm (cf. Viola & Jones 2001) providing a rectangle at image coordinates from the camera image. Persons with their face turned away from the robot are disregarded. Thereby, it can be assured that only those are attended to which show signals of attention themselves.

For a combination of proximity-based input and attentive behavior, first of all the distance between human and robot has to be estimated. For the purpose of the system, it is sufficient to perform a triangulation of the detected facial area with inherent camera parameters, i.e. the opening angle and image size. As to *prompt* attendance, the robot gazes in the direction of the human, which is accomplished when the person's face is detected in the center of the camera image. The deviation of the current viewing direction of the robot from the target can therefore be incorporated as an parameterizable output. It can be obtained by comparing the facial location with the center of the in-eye camera image.

Figure 5.1 gives an idea on how the robot's attention is further modulated by the distance to a visitor with the help of two independent principles. As a first measure, compensation angles Φ_{pan} and Φ_{tilt} are introduced that are heavily dependent on proximity parameters. The robot constantly turns by these angles in order to keep the person's face in the image center. Their values are determined by the width and height normed vertical (d_y) or horizontal (d_x) deviation of the face from the image center multiplied with a basic angle ϕ. Regarding the individual *proxemics* zones, the following basic angles are incorporated: Intimate distance employs a basic angle of $\phi = 2°$, $\phi = 1.5°$ is used for personal, $\phi = 1°$ for social, and $\phi = 0.5°$ for public distance. If the angle is below a threshold ϵ no movement is performed:

$$\Phi_{pan} = \begin{cases} -\phi & d_x > \epsilon \\ \phi & d_x < -\epsilon \\ 0 & \text{otherwise} \end{cases} \qquad \Phi_{tilt} = \begin{cases} -\phi & d_y > \epsilon \\ \phi & d_y < -\epsilon \\ 0 & \text{otherwise} \end{cases}$$

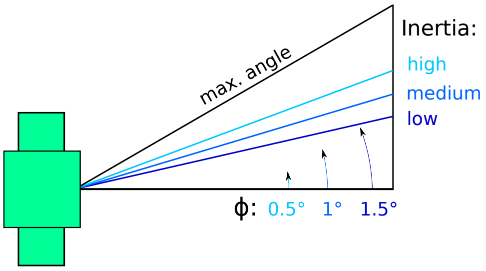


Figure 5.1.: Distance dependent modification of robot movements in a single joint as implemented for the video study. Values of the basic angle ϕ used for compensation and inertia settings with regard to distance class are depicted. Dark blue marks the values used in personal distance, values in the social distance are highlighted with a lighter blue, and values with the lightest blue are used in public distance.

The resulting turn angle Φ_{pan} is distributed among the hip, head, and eye joints. The head and eye joints combine to the overall pitch angle Φ_{tilt}. Because the resulting compensations for the 2d deviation in the image is already distance specific, this method leads to a engagement of the robot which becomes stronger while the person comes closer.

As a second method incorporating social distance into the attention mechanism, the relative compensation angles are decomposed into robot postures based on social distance. With a distribution of the angle on different joints depending on the distance class, the usage of certain joints can be restricted or emphasized (cf. Fig. 5.1). The robot for example appears more stiff, if the head movements are restricted or more vivid if it contributes to the overall gaze direction. Furthermore, a vis-a-vis formation can only be established reliably if the hip is not held back. For this purpose, a so-called inertia value (in the sense of stiffness) determines to what extend the complete range of a joint is being exhausted. A virtual boundary limits the theoretically possible angle that a joint can be maximally moved.

With a high inertia value the individual joints are limited least, i.e. they can be moved to 50% of their real maximum. Because of that, most of the movement is accomplished using the eyes only. The head is used for changes in gaze directions that cannot be reached by the eyes alone. The hip remains practically unused. When the inertia is set to medium, the joints are virtually limited to use only 40% of their range. In this setup, the head is used much more frequently for changing the posture. A low inertia value limits the joints to 30%. Therefore, also the hip joint contributes very often to the actual turn value. The limitation above does not introduce a hard boundary, but a soft one instead. If the angle cannot be distributed the aforementioned way, then the remaining part will be added to joints that have not already reached their real maximum.

Since according to Kendon (1967) humans do not stare consistently at each other during a conversation, the implementation of distracting random gazes is suggested as a reasonable addition to the attention mechanism. These shift the robot's focus from a human to another location for a short time of approximately one second. The robot's attention seemingly gets caught by some other entity in the room. During the video study, the current view angle is shifted relatively from the current gaze location and is decomposed exactly the same as in the case of a detected face. The only difference is in the usage of joints. The inertia value is even higher than if a human is detected. Thus, the joints are only limited to 70% of their range. This way, one can assure that the robot does not turn its body away from a human in a face-to-face situation.

5.2.2. Evaluation of Attentive Strategies

The feasibility of the spatial attention strategies during approaching has been evaluated with the help of an on-line questionnaire (cf. Appx. E). Participants had to answer questions referring to videos that show a human approaching the receptionist. Further they had to mark the time of the robot's first interaction attempt in the videos. Two main questions have been addressed in this survey:

1. To what extent does the dynamic modification of the attention behavior alter people's perception of the robot?

2. Which influence does the addition of random gazes have on the perception of the robot?

For the questionnaire the beginning of an interaction between human and robot has been videotaped for each condition. This way, it can be ensured that each participant group rates exactly the same robot behaviors. Furthermore, the experimental results can not be influenced by the various ways people would try to interact with the robot. Comparability within and between participant groups can only be guaranteed because the interacting person's behavior, especially his path towards the robot, stays the same in all videos.

The following common procedure has been presented in each video: A human enters a room, walks through it, and eventually stands in front of a desk with the robotic setup. When the human arrives and enters the robot's personal distance, it says: "Hello, my name is Flobi. How may I serve you?". The human answers: "Tell me the way to Patrick's office".

To evaluate the distance-adapted attention model, dynamic movement styles are compared to two static behaviors. If the robot behaves dynamically, inertia value i as well as compensation angles Φ_{pan} and Φ_{tilt} are adapted to the actual distance of a person as introduced in Section 5.2.1. Contrarily, static behavior in close and far interaction styles implies that while the human is approaching, these parameters are fixed. The robot behaves as if an interlocutor would be located in either a personal (close) or public (far) distance to the robot. Furthermore, to also investigate the influence of in-between random gazes, it is differentiated among normal movement styles to normal plus additional random movement. As a consequence, eight videos of the same situation but with different interaction styles have been recorded:

Z	The robot does not move at all, resulting in [Z]ero movement.
R	The robot's gaze is shifted only [R]andomly.
CN	The robot tries to focus its counterpart but acts as if he were permanently in a [C]lose distance, [N]o random movements added.
DN	The human is focused. This time, the movement is [D]ynamically adjusted to the distance.
FN	The gaze is shifted as if the person were in [F]ar distance.
CR	Same as CN, but [R]andom movements are added in between.

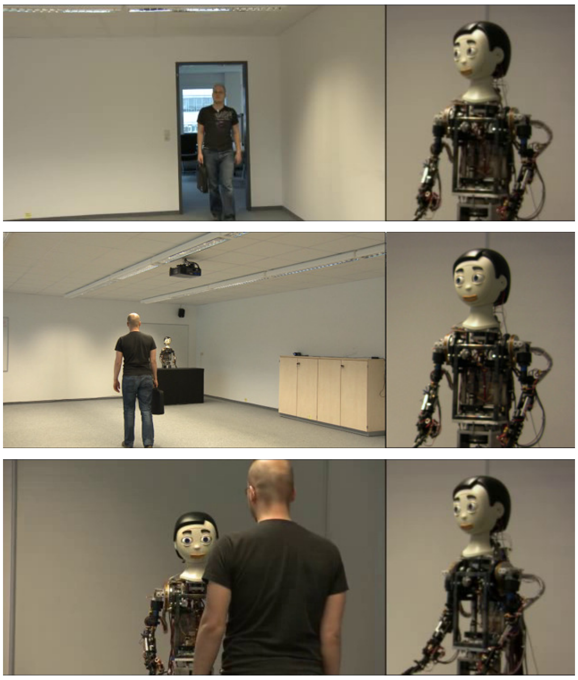


Figure 5.2.: Video screen-shots from the approaching study. The left camera image follows the person as he comes closer to the robot. In the right image a close-up of the robot is shown to let people identify the robot's motions reliably.

DR Distance dependent as DN, but with random movements.

FR Like FN, with additional random attention shifts.

The interaction has been recorded from two perspectives. One camera has been following the human all the time and another one shot a close-up of the robot. Both of the videos have been combined to a single one that shows the perspectives side by side. In Fig. 5.2, three screen shots of the resulting video that has been shown to the participants are presented. All of the videos have been synchronized to the frame one could spot the robot in the left video for the first time. They fade to black while the human answers the robot to suggest an ongoing interaction between the two agents.

Participants had to fill out a questionnaire where they have been shown three different videos. The first video always showed the Z condition, in the second and third video, the participants could see two videos from different conditions. To prevent side effects of sequence, these videos have been shown in random order. Altogether, participants have been put in one of the following five experimental conditions:

NR Videos differ in containing [R]andom movements or [N]ot. (DN and DR, or FN and FR, or CN and CR)

FD The robot acts as if the human is either [F]ar away or dynamically adjusts its movement to the [D]istance. (FN and DN, or FR and DR)

CD The robot treats the human either as [C]lose to the robot or dynamically adjusts to the [D]istance. (CN and DN, or CR and DR)

CF The robot acts as if the human is either [C]lose or [F]ar away. (CN and FN, or CR and FR)

RR The robot only shows [R]andom movements in both videos. (Control group)

For each of the videos, participants had to determine the *Timestamps* when they thought the robot had realized that the human wanted to interact with it. They had to do so by stopping the video at exactly this time. The video could not be watched any further beyond that point. After identifying the timestamps in all three conditions, the videos have been presented a second time. Here, participants had the possibility to watch the video as a whole and as many times as they wanted. Beneath the video, they have been asked to rate certain aspects of the robot's behavior on a five-point scale (0-4) (cf. Likert 1932):

- The robot's *Interest* in the human
- The *Appropriateness* of the robot's behaviors
- The movement's *Human-Likeness*
- The *Naturalness* of the robot's movements

- How much *Attention* the robot payed to the human.
- The robot's *Autonomy*
- How much of its *Intention* the robot revealed.

Altogether 111 users have participated in the study, of which 39.6% are female and 60.4% are male. Their age varied between 16 and 70 years with an average of 30.5. Almost half of them are affiliated with the university at participation time, either as students (31.8%) or as scientific staff (18.2%). The vast majority of 88.3% consists of native German speakers. The rest states a high understanding of English or the German language. The questionnaire is available in English and German languages, so the questions could be well understood and answered by every participant.

The robot experience highly varied between subjects. A very large part (84.7%) rates their robot experience lower than average on a five-point scale (0-4). The mean value for the participant's robot experience lies at around 1.04. In contrast, most of the participants rate their own computer experience either 3 or 4 (67.9%). With an average of 2.94, the computer knowledge seems to be fairly high among the participants. In general, one can say that although the majority of participants are naïve to the subject, they apparently have a common technical understanding.

Pausing time of the video and answers to the questionnaire have been evaluated for significant deviations of their mean value. As a method for the comparison, a paired-samples Wilcoxon signed-rank test with a significance level $\alpha = 5\%$ is used.

Goal Directed Movements Almost all of the posed questions result in significantly different ratings between the Z video (zero movement) and every other video that has been shown. Participants rate all of the robot's attributes higher for videos that showed a moving than for a still robot ($p < .037$). Also, participants assume the robot realizes its human interaction partner faster if it was moving. Response times in the stopping task are significantly shorter compared to the no-movement condition ($p < .009$).

The [RR] group with 12 participants is an exception to the others: Fig. 5.3 shows in detail that videos containing pure random movements only produces significant changes in the participants' ratings for the robot's *Human-Likeness* and *Attention*. Instead, *Interest*, *Appropriateness*, *Naturalness*,

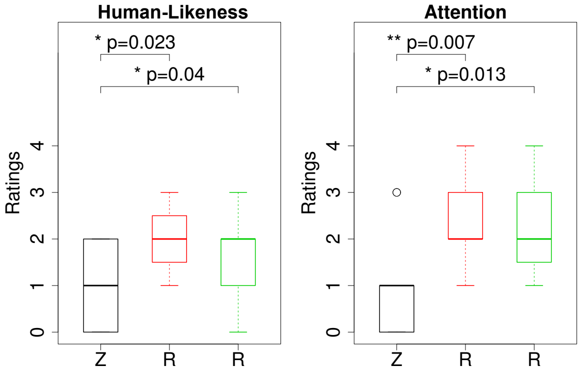


Figure 5.3.: Box-plot of ratings in the [RR] group. The zero movement condition [Z] is compared to two different random only movement types [R]. Median values are marked with a bold line, the box contains central 50% of given answers. The rounded (3 dig.) two-tailed significance p of the statistical test is depicted if the differences of means are either significant ($*$) or highly significant ($**$).

Autonomy, and *Intention* can not be distinguished from videos without any robot movement. Only the first of both random videos has been stopped significantly earlier than the video without movement ($p < .024$). Pausing times of the second random video are higher again and hence no significant differences can be detected. In Fig. 5.4 the probability density functions calculated from the video timestamps for the RR group are shown. There is an obvious difference between the densities for the video without robot movement [Z] in comparison to both random videos [R]. While the former consists of a single peak at the end of the video, densities in the latter are more distributed in time and compound of two flatter maximum points. Additionally, a minor shift to the right for the second random video is noticeable.

Distance Dependent Modification of Behaviors Only one of the FD, CD, and FC groups show significant deviations in the ratings of the robot's behaviors. Groups CD (21 users) and FC (24) do not show any differences between the two videos that have been presented to the participants. Re-

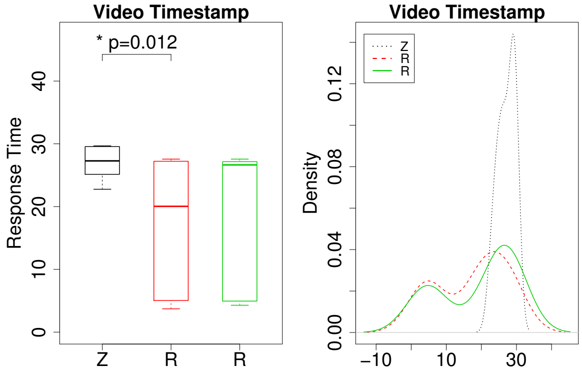


Figure 5.4.: Box-plot and density of the video response time in seconds from the random-only [RR] group. The zero movement condition [Z] is compared to two different random only movement types [R]. Median values are marked with a bold line, the box contains central 50% of given answers. The rounded (3 dig.) two-tailed significance p of the statistical test is depicted if the differences of means are either significant ($*$) or highly significant ($**$). Densities for [Z] as well as for both [R] videos are exemplary given to the right.

sponses in the FD condition (26 participants, FR vs. DR or FN vs. DN) instead can be distinguished. The result of this comparison is shown in Fig. 5.5. The robot's initiative is spotted earlier and participants rate the robot's *Interest*, *Attention*, and its *Intention* higher in the video showing the distance dependent behavior than in the far away condition.

The Influence of Random Movements The participants' answers of the NR group (27) differ significantly in five categories. Please refer to Fig. 5.6 for detailed results. The robot's *Interest*, *Human-Likeness*, *Attention*, and *Intention* is rated better in videos with random movements (CR, DR, FR) than in videos without random movements (CN, DN, FN). Also, the robot's intention to communicate has been perceived earlier if in-between random movements occur. Other attributes do not show significant differences in the users' ratings.

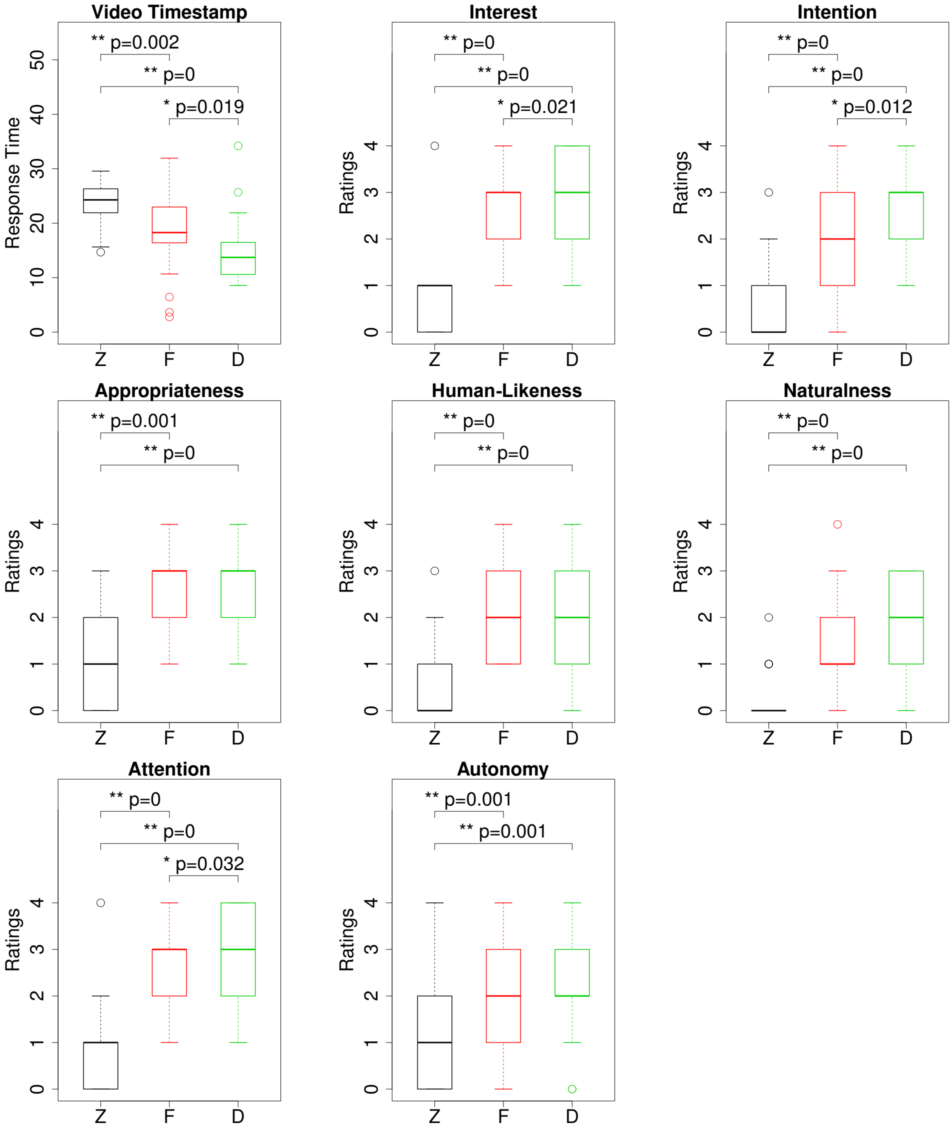


Figure 5.5.: Box-plot of video stop time in seconds (cont.) and ratings (disc.) by video type of the [FD] group. [Z] marks zero movement videos, [F] consists of videos from the far away condition, and [D] contains videos with dynamic movement adaption. Median values are marked with a bold line, the box contains central 50% of given answers. The rounded (3 dig.) two-tailed significance p of the statistical test is depicted if the differences in the ratings are either significant ($*$) or highly significant ($**$).

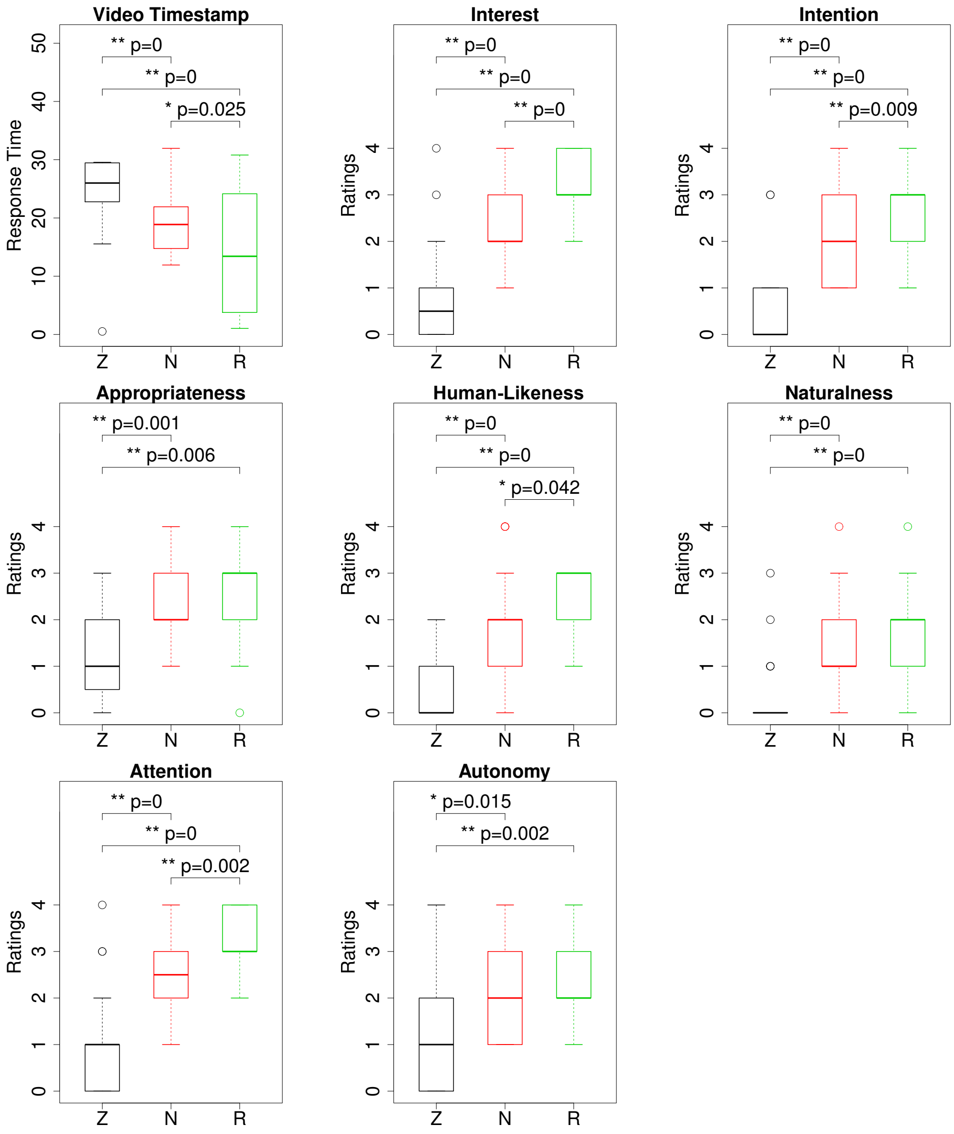


Figure 5.6.: Box-plot of video stop time in seconds (cont.) and ratings (disc.) by video type of the [NR] group. [Z] marks zero movement videos, [N] consists of videos with straight person-directed gaze, and [R] contains videos with additional random gazes. Median values are marked with a bold line, the box contains 50% of given answers. The rounded (3 dig.) two-tailed significance p of the statistical test is depicted if the differences in the ratings are either significant ($*$) or highly significant ($**$).

5.2.3. Implications for the Approaching Phase

The above results demonstrate that the presented strategies can serve as an encouragement and entry method for face-to-face interactions between human and robot. Each of the presented movement types is more appealing to a human user than no movement at all. They however also indicate that an embedding into a larger repertoire of behaviors in an applied scenario might be necessary to strengthen the approach.

Even totally random movements (RR group) suggest a certain human-likeness of the robot. The significance in the ratings of the attention in the random-only case might be caused by the fact that the robot accidentally looked straight into the human eye as it began to speak. If this had not been the case, the attention ratings of the random behavior would possibly also not be distinguishable from the no-movement case. Another possibility would be that participants attribute the robot some kind of attention because it can shift its gaze to places somewhere in the room.

On the one hand, participants identified an initiative by the robot earlier in the first random video compared to no movement. On the other hand, in the second random video, the timestamps could instead not be distinguished from the zero condition. Participants apparently mis-interpreted the random movements as a sign of interaction in the first place, but realized the movements are intention-less while watching the second video. Random-only gaze-shifts therefore are not sufficient for leading into an interaction but nonetheless can be employed as an idle behavior signaling a general readiness of the robot.

In contrast, random gazes in conjunction with person-directed gaze can lead to a better user experience than person-directed gaze alone (NR group). Participants believe that the robot has more interest in the human, is more human-like, pays more attention to the human, and expresses its intentions to a greater degree when the robot additionally exhibits random gazes. Also, they noticed a robot-triggered interaction earlier in this case.

At a first glance it might be confusing that especially the attention is rated higher when the robot looks away from time to time. An explanation could be that these distracting gazes actually help to communicate an attention mechanism to the human because the robot re-focuses on the human every time it had looked away. Therefore, the robot shows that its attention is caught again and again by the human. As a result, the robot communicates that it is interactive in an effective way that can easily and almost immediately be detected by a human interaction partner. Additionally, Kendon (1990, Chap. 6) describes a mutual approaching between humans that involves disengaging by looking away from the other followed by re-focusing, which explains the increased human-likeness of the behavior. While obviously random gazes help to assign a certain personality to the robot, they at the same time do not have a negative influence on the appropriateness and naturalness of the behaviors or the autonomy of the robot. As a consequence, the robot apparently does not lose any of its functionality by the addition of distracting gazes.

No significant differences can be found in the ratings and timestamps between the groups that contain the two distance independent behaviors of the robot (FC group). The actually different behavior styles in these conditions apparently do not lead to a higher valuation in one of them. While all cases in this group differ significantly from the zero movement video, participants do not prefer one solution over the other.

Also, the distance-dependent condition is not distinguishable from the condition in which the robot acts as if the person stands directly in front of it (CD group). It is believed that this could be caused by the similarity of the videos for these cases. Participants are not able to tell the difference between the two conditions. That might be a problem of the video itself but could also be a consequence of the experimental setup. Since people are not in the same room with the robot but only rate a video instead, their comfortable feeling could not be violated by a robot that doesn't respect personal distances. Therefore, the ratings for the robot are almost identical in the case of direct response as in the dynamic case.

Between the far-away and the distance-dependent condition (FD group), significant differences can be spotted in the user's ratings of the robot's interest, attention, intention, and video timestamps. Apparently, the robot is experienced as more responsive and expressive in general, if it uses more of its capabilities and turns its body earlier and more frequently to the interaction partner. As these movements are perceived sooner and rated higher, the distance-dependent behaviors should be preferred over the artificially restricted ones.

5.2.4. Conclusion

The video study demonstrates that a spatially aware attention mechanism can serve as an entry point for a face-to-face interaction in a receptionist scenario and should be preferred over other strategies such as a non-moving or only randomly turning robot. As a consequence, a step-by-step signaling of attention is proposed in order to transition from distant *HRI* to a close interaction in the approaching phase.

While random movements alone are not suitable as an entry for the interaction, the overall behavior can benefit from the addition of random directions to the person-directed gaze in terms of user experienced robot intention, attention, interest and human-likeness. Involvement of the robot should be shown in a distance dependent manner to increase the perceived intention, attention and interest. Restricting the robot's hip movement in face-to-face situations leads to a lower overall rating of the robot's responsiveness. The opposite case of immediate response remains to be evaluated in greater detail (cf. Chap. 6), since it has not been possible to produce any significant differences in the video study. It nevertheless remains expected that an immediate turning response would not be appropriate under real-world conditions given the related literature.

5.3. Mixed-Initiative Opening

A pro-active robot greeting is proposed as an appropriate strategy for the receptionist to open up the dialog phase. It is supposed that humans can be led into the dialog by the robot's incentive if they reach a convenient interaction distance as well as configuration in terms of proper formation and adequate attention towards the robot. Furthermore, it is believed that as a direct consequence of steadily increasing robot attention that has been discussed in the previous section humans to some degree expect a verbal robot utterance.

Kendon (1990, Pg.192) observes that a close salutation between two people usually involves a brief halting, with their bodies directed towards each other. It is therefore suggested for the robot to face an approaching interaction partner during dialog opening with the whole body including eyes, head, and torso. In contrast to humans, the receptionist is fixed at hip level so that it cannot alter the orientation of its feet. This a fact however, can only be regarded as a minor limitation, as the feet are not visible for the human in close interactions with the robot. The upper body still is able to turn towards the human counterpart leading to an *F-Formation* in a vis-a-vis orientation.

It is proposed to perform a robot initiated greeting only if the imminent interlocutor faces the robot and thereby enters the closer phase of a social-consultive interaction distance between adults according to Hall (1966, Chap. X). Within such circumstances, one can proceed on the assumption that the human is willing to communicate with the robot face-to-face as she is actively approaching the robot while expressing attention towards it. The specific distance is suggested on the fact that the social-consultive distance corresponds to the farthest communication distance used among humans and therefore qualifies for an interaction opening. Nonetheless, as the robot is somewhat smaller than the human plus the situation remains new and unexplored to the person, the close phase is chosen as the required minimum distance for a receptionist robot to actively open an interaction by itself.

Besides the robot-initiated opening, a possibility for the human to engage in the dialog oneself at any time during approaching is suggested. Thus, also human efforts are taken into account as means to open the dialog. Otherwise, in the case of simply disregarding such signals, obviously the person's confidence in her actions during the interaction is severely lowered, which most likely leads to a worse user experience (cf. Chap. 2.1.3). Because of that, a mixed-initiative dialog system for the receptionist is needed allowing for openings triggered by the human as well as for those triggered by the robot caused by the current spatial configuration.

5.4. An Integrated Concept Leading into Dialog

This section presents the integrated concept for leading into a dialog between human and robot as a consequence emerging from the aforementioned strategies. It covers the proposed phases and signals until the mutual dialog has been successfully established (cf. Chap. 3.2) and therefore addresses Claims I, II, and III. The complete concept is incorporated into the spatially aware receptionist and meant as an alternative opening leading to the basic interaction that has been introduced in Chapter 4.3.

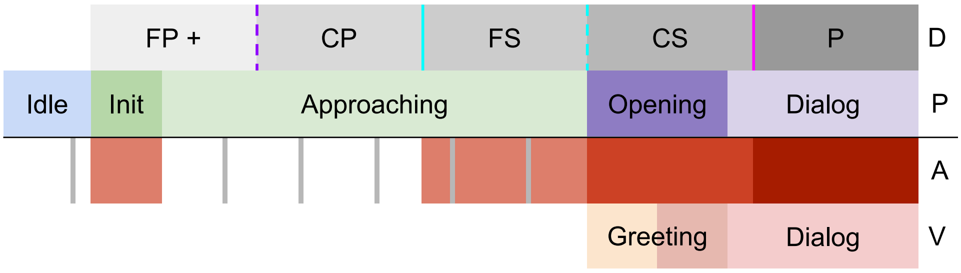


Figure 5.7.: Schematic depiction of the integrated spatial concept leading into the dialog phase in the advanced receptionist. At the top row, the distance class [D] between robot and human is given, namely far public [FP] or farther, close public [CP], far social [FS], close social [CS] and personal distance [P]. The second row describes the current phase or signal [P] of the encounter. The bottom rows describe robot behavior, i.e. rising attention towards the human in red interrupted by random gazes (gray) in [A]. Verbal utterances [V] below describe a robot initiated greeting followed by a human answer and a shared dialog.

Figure 5.7 schematically depicts the here suggested behavioral enhancements. During the idle phase, the robot regularly changes the direction of its gaze towards a randomly selected target in order to signal availability even if no possible interaction partner has yet been detected. Such random gazes are depicted as gray segments in the robot's attention [A] of Figure 5.7. Exactly the same way as in the basic receptionist, they are continuously applied during approaching until the interaction is opened. During dialog, arguably gaze does not need to be artificially distracted as attention has to be divided anyways between human, floor plan, and the location in question.

On recognizing a person, the robot sends out an initiation signal as discussed in Chapter 5.1 with the intention to lead into an approaching behavior. Such a signal is sent instantly, i.e. at a far public distance according to Hall (1966, Chap. X) if the facilities the robot is residing are of appropriate dimensions. The signal consists of a short gaze towards the human that immediately turns back into the previous direction and only employs head and eye movement. As such a gaze implies a form of attention towards the human, it is colored in red in part [A] of Figure 5.7.

A distance dependent attention system is incorporated into the advanced receptionist robot during approaching as a consequence of the user study presented in Chapter 5.2. It has been slightly adapted based on the findings of the user study in order to better suit the current robot as a platform and also to better integrate into the overall concept. As a part of that, the attention system has been altered in contrast to the one evaluated in the study. The artificial restriction of joints has been removed in favor of a more realistic gaze control mechanism by Pattacini (2011) which implements human-like head-eye coordination according to Guitton & Volle (1987).

As a consequence, no further actions besides random gaze shifts are performed by the robot until the person decides to enter the social distance. Upon arriving in the social distance the robot begins to focus its possible interaction partner with head and eyes similar to the initiation signal. Attention directed towards the human is colored in shades of red in Figure 5.7.

Only at the close phase of the social distance, also the hip is being integrated to turn the torso slightly (50%) towards the person creating a new spatial configuration for opening up the interaction. As examined in Chapter 5.3, the robot uses a verbal utterance to lead over into the dialog phase. More specifically, the statement is “Hello, my name is iCub. How can I help?”. Figure 5.7 [V] gives an idea of how this secondary opening is employed as an additional method besides the traditional user-initiated opening as depicted in Figure 4.4. Instead of a need for the human to act first, the interaction is opened by the robot based on the distance between the two, followed by a human answer that finalizes the transition into the dialog phase.

Because the floor plan of the receptionist is involved in the dialog, the human usually enters the far phase of the personal distance during opening or shortly after the dialog begins. Such a distance is motivated due to a more comfortable handling of the map as well as the establishment of an vis-a-vis *F-Formation* resulting in a common *interaction space* with the robot. In order to maintain such a formation robot-wise, the robot uses its full potential of hip, head, and eyes for displaying attention towards the human in personal distance. In Figure 5.7 [A], darker shades of red accordingly imply an increased amount of torso-movement that contributes to the emitted robot attention. As a further corollary of distance awareness, the receptionist requests its interlocutor to recede, if detected inside the intimate surroundings.

Such an integrated opening greatly supports Hypothesis H. Already at far distances, the robot is able to signal availability to a possible interaction partner with a short gaze (Claim I). An attention mechanism based on spatial configurations allows the robot to gradually indicate interest in helping the human interlocutor (Claim II). It is therefore proposed that the robot is able to lead into a dialog and offer services by itself on detecting a human in its social interaction area. Distance and attention initiated greetings enable the robot lead into the dialog pro-actively which opens up an alternative method that reduces uncertainty (Claim III). Concepts presented here will be evaluated alongside the upcoming strategies in Chapter 6.

5.5. Social Gesturing during Dialog

In this section, enhancement strategies for the dialog phase between human and robot are presented that aim at supporting Claim IV. It presents additional means of communication that are supported by a spatial model and applied supplementary to the standard dialog as presented in Chapter 4.3.

For spatial awareness in face-to-face interaction, it is proposed for the robot to maintain a model of human activity in its close surroundings called the active peripersonal space as introduced in Chapter 3.4. With the help of such a model, the receptionist can make assumptions on where to point at (map or real world), which hand to use, and whether there is the need to use *spatial prompting* (cf. Chap. 2.2.4) for reaching a target. Additionally, its spatial awareness enables the robot to modify the duration of the gesture and whether to use gaze or not.

In order to reflect the attributed right-handedness of the robot, its right hand is employed as a default when conducting a gesture towards the map or towards a real world location. Only in two cases, the left hand is used as a fall-back device. Either due to obstructions by the human that are inferred from the active peripersonal space given that the left arm provides a more unimpeded access. A second reason for choosing the fall-back emerges if the target is difficult or impossible to reach for the right hand as opposed to the left one. In practice, such a strategy leads to the left-most rooms on the floor plan to be pointed at with the left hand, while all others are addressed with the right hand if there is no human activity.

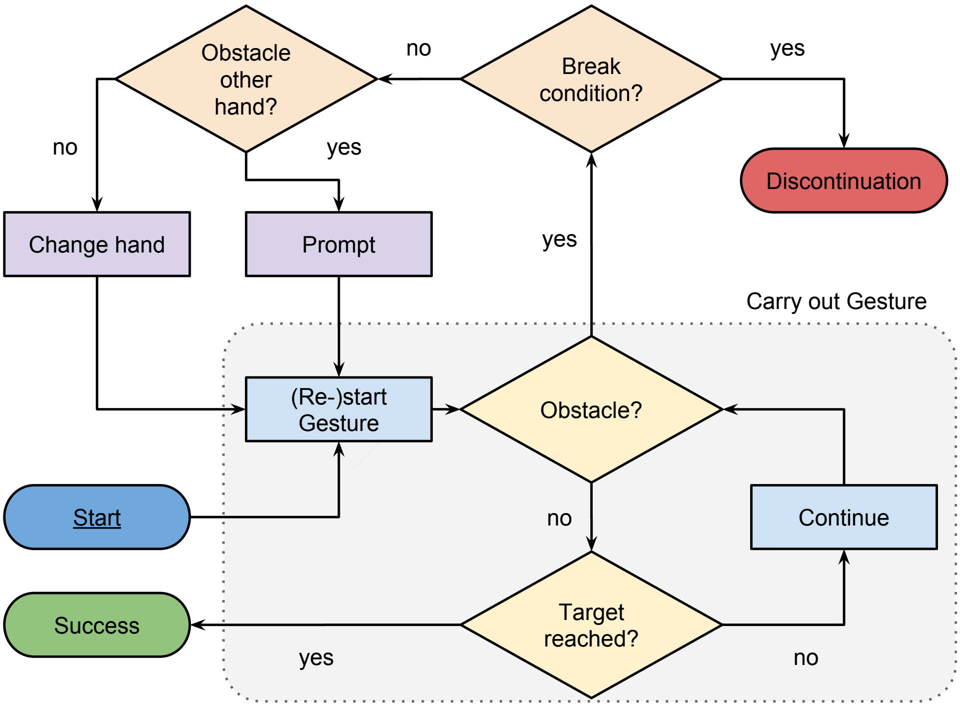


Figure 5.8.: Flowchart of the pointing gesture in the advanced receptionist. In contrast to the basic receptionist (cf. Fig. 4.5) depicted in yellow and blue shades, the advanced procedure (orange and purple shades) involves the changing of hands and a *prompting* mechanism to reach the pointing target. The gesture is only aborted if a defined exit condition is fulfilled.

As to enable a distance dependent awareness of human presence during the dialog, the active peripersonal space is employed to resolve possible obstructions during pointing gestures by the robot. Such an awareness is considered to have several advantages over traditional methods that as of now mostly consist of error handling, i.e. cancellation of the gesture. It is assumed that spatial prompting can be used as an appropriate method to induce a human to retreat their hands from areas of interest for the robot. It seems to constitute an acceptable comprise between discontinuation or immediate abandonment of the gesture on the one hand and carrying on regardless on the other hand. Spatial prompting is therefore expected to solve territorial conflicts between human and robot without the need to cancel a task, while leading to a better understanding of robot capabilities and demands in the human. The dialog with the receptionist should as a result profit in terms of usability and security.

Figure 5.8 proposes under what circumstances a spatial prompting mechanism can be embedded in a gesturing system. It furthermore describes how spatial prompting is realized during the dialog phase to enhance the standard gesture in the receptionist robot. As long as a pointing gesture is carried out, the basic receptionist system monitors the active peripersonal space for obstacles that are caused by human presence. If an obstacle blocks the way towards the desired target or the target itself, in contrast to the basic receptionist (cf. Chap. 4.3) the gesture is not simply discontinued with an error message as depicted in Figure 4.5. The advanced receptionist instead, explores more sophisticated strategies that still aim to reach the target in a socially aware manner.

At first, the system checks whether the target might be reachable through an alternate route, using the other hand. If that is the case, the current gesture is stopped and then restarted on the second arm. Otherwise, a prompting strategy is employed in order to signal a pointing intent towards the location and therefore motivate the human to recede from the area of interest. The approach in this thesis constitutes of repeated spatial prompts towards the target in conjunction with short utterances as long as the target is being occupied by the human. Thereby, the extend of the gesture is steadily increased with the aim of simultaneously enhancing the distinctness and eventually the humans awareness of the robot's intent.

A prompt is implemented as a multi-modal *spatial signal* in the advanced receptionist. It consists of a short utterance of one or two words in conjunction with a reaching gesture and a short gaze towards the target area. Thereby, the robot's hand is opened with the fingers adjacently directed towards the target and the palm oriented towards the floor. The movement stops at a predetermined percentage of the distance, leaving the human time to recognize the robot's intention and react accordingly. At halting, visual attention is again shifted towards the human in order to signal an expected retreat. The robot then carries on with the obstacle detection routine and redirects attention to the location in question. If the pathway is still blocked, the prompting mechanism is repeated with altered parameters. The prompting utterance varies in the order “This...”, “This is...”, “Sorry...”, whereas the hand at first moves ten percent of the distance towards the target location. Subsequently, the gesture’s extent is increased by another ten percent each trial. If the prompting does not result in a free access after three attempts or times out otherwise, the gesture is eventually discontinued. In that case, the robot states “Sorry, you’re in my way.” as a notification of cancellation (cf. Appx. C).

As a further consequence resulting from the social properties of the peripersonal space, territoriality is regarded in a distance dependent manner. To begin with, the explicitness of prompting already is modulated by the target’s distance from the robot as it is designed as a function of distance (ten percent). As a second measure, prompting behavior is prohibited in order to better respect occupancy effects (cf. Chap. 3.4.3) if the obstruction is within the close proximity of the robot. The robot then acknowledges the human presence as an extraordinary effort (cf. Chap. 2.2.3) and discontinues the pointing gesture.

As a summary, this section introduces spatial prompting as a strategy to reduce uncertainty during the dialog phase between a human and the receptionist robot. It furthermore explains at which occurrences during the dialog such a behavior can be incorporated into the existing system. It therefore delivers an applied concept in order to address Claim IV. The behaviors integrate well into the earlier presented strategies as they can be operated in parallel to attention mechanisms and can be easily integrated into the overall scenario.

5.6. Closing an Interaction and Departing

In this section, a sophisticated method for closing an interaction between human and robot that smoothly leads into disengagement according to Chapter 3.3 is proposed. On the one hand, presented methods in this chapter are therefore aimed to fit into the overall approach of spatial awareness and on the other hand they are conceptualized with regard to individual concepts that have already been introduced. After refining their immediate purpose in context of the system, it is described at which point each strategy is applied to the basic receptionist described in Chapter 4.3. All of them, however, are situated at the end of the dialog and afterwards.

A first concern at the end of the dialog emerges from previous interactions with the receptionist. For a great portion of naïve users it does not seem to be clear at which time an interaction is over, although clearly stated in the instructions. From observations, it appears as if they are either not aware that they are allowed to end the dialog themselves by stating farewell words, or still expect the robot to execute some action. Both explanations however have in common that users have an inherent uncertainty about the situation. Such circumstances not only badly influence the general user experience but also imply unforeseeable effects on user ratings regarding the interaction in general. To overcome such a side-effects and therefore address Claim V, it is proposed to incorporate an alternative dialog closing that is initiated by the robot. Similar as in opening strategies (cf. Chap. 5.3), pro-active closing seems to qualify as an appropriate and flexible method if employed as a secondary option besides user-triggered closing.

In the advanced receptionist, such behavior is incorporated into the default dialog between human and robot at the table as described in Chapter 4.3. Robot-initiated closing is thereby employed as a complementary method besides the traditional user-initiated dialog closing as depicted in Figure 4.4. It is emitted only under the following circumstances: Firstly, as the robot always asks if the user needs additional help, it closes the dialog if this question is negated. Secondly, if there are no further questions after a certain time has exceeded, the dialog is closed as well by the receptionist.

Furthermore, as reported by Ghosh & Kuzuoka (2014), changes in body posture, more particularly leaning backwards, can act as an accelerator if a robot tries to signal the ending of an interaction. Coherently to Claim VI, it is therefore considered beneficial for the robot to also emit such a supplemental signal at dialog closing in order to induce departing behaviors. Besides that, it is proposed to apply the same spatial attention mechanism during departing that has been described for approaching in Chapter 5.2. As an direct consequence, the robot is able to continuously signal availability in a descending manner and thereby enable a seamless continuation in the case the human plans to resume the dialog.

Therefore, the advanced receptionist supports a disengagement by leaning backwards during the farewell utterance and thereby lowering its head so that no further attention is signaled towards the human. After a short timeout of seven seconds, the attention mechanism is re-enabled and the robot turns towards the human if she is attending the robot.

In summary, a number of *spatial signals* are proposed as well as incorporated into the advanced receptionist in order to support Hypothesis H during dialog closing and departing. At first, robot-initiated dialog closing is employed as an alternative to the usual user-triggered course of action. Such a strategy is accompanied by a disengagement signal in order to lead into departing and thereby address Claim V. During departing Claim VI is approached. Attention towards the human is shown in a shrinking amount that depends on distance to the robot so that the interaction is smoothly fading out leaving open the possibility for resumption. These spatial strategies are evaluated alongside every other measure applied to the advanced receptionist in Chapter 6.

5.7. Summary

The proposed strategies in this chapter are an example on how to enhance a robot with social interaction strategies in the domain of spatial awareness for a better user experience during *HRI*. The chapter furthermore provides a precise overview how the strategies are embedded in the basic system's course of action as given in Chapter 4.

As an illustration how the basic system has been altered in order to enable the advanced receptionist, Figure 5.9 consists of a renewed version of Figure 4.3 displaying the software components of the system. Especially the situation model is a lot more elaborate, taking social parameters such as proximity and spatial configurations as well as presence and activity in the peripersonal space into account. Still, the dialog component plays a major role in the software by deciding when to conduct an action. However, it is adapted and modulated by a spatial awareness concept that is able to exhibit socially adapted behaviors. Thereby, the output possibilities, especially the display of attention and salutation routines can act in a more sophisticated fashion.

At first, this chapter has introduced the idea of a short gaze towards the human in order to initiate an interaction in Section 5.1. The following approaching phase has been investigated with the help a prototype implementation displaying attention in a distance dependent fashion. Based on results of a conducted pre-study, this behavior has been proposed as feasible during approaching in Section 5.2. Next, in Section 5.3, a mixed-initiative

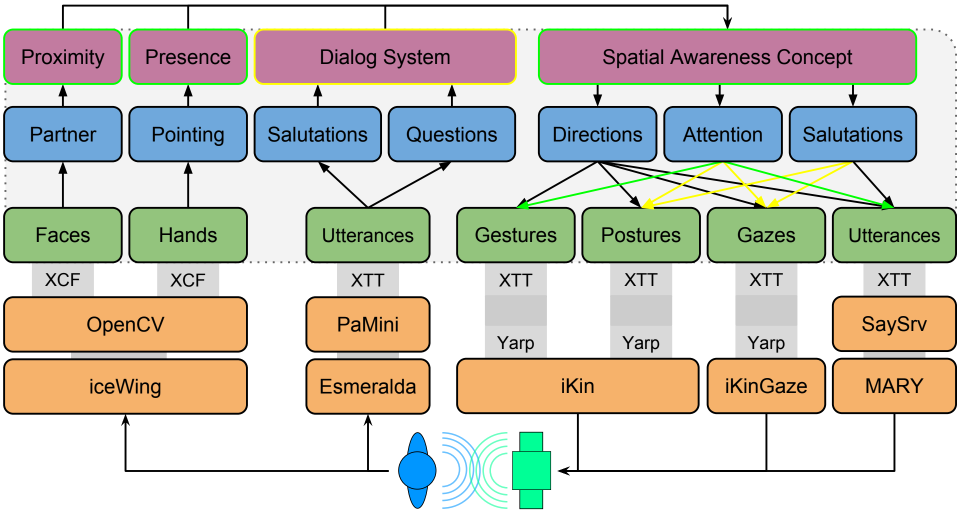


Figure 5.9.: Spatially aware way of proceeding in the advanced receptionist. An addition or adaption in contrast to the basic system (cf. Fig. 4.3) is marked in green or yellow. The topmost layer additionally includes models about social configurations, namely proximity and presence that both have an influence on the dialog. As a consequence, the complexity of robot behaviors is altered in order to cover a higher number of modalities.

dialog concept has been proposed for opening an interaction based on social distance and configuration. An integrated concept for initiation, approaching, and opening is then introduced and related on the implementation of the basic receptionist in Section 5.4.

Afterwards, in Section 5.5 social presence and activity are incorporated into the receptionist as sources for enhancing gestures during the dialog phase. As a result, the section proposes *spatial prompting* strategies to be added to the standard gesture in order to increase the probability that the robot is able to perform its gesture while not irritating the user. Contrary, prompting is aimed at making the interaction more stable. The chapter furthermore gives details on how prompting is implemented in the basic receptionist.

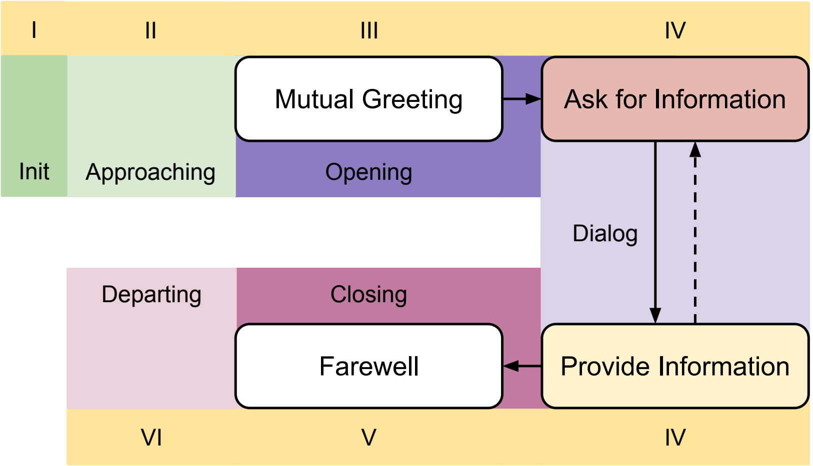


Figure 5.10.: Sequence of a dialog between the receptionist and a human visitor in the advanced receptionist. In contrast to the basic dialog as depicted in Figure 4.4, it is embedded into a complete situation. During each phase, the receptionist reflects one of the claims in its behaviors to address Hypothesis H.

Section 5.6 is finally concerned with appropriate means to end the dialog in a spatially aware way. It again proposes self-initiative as an alternative for dialog closing and employs the active lean-back signal in order to induce departing behavior. Additionally, it proposes the spatial attention mechanism to be also incorporated during the departing phase as to enable a seamless resumption of dialog.

As a result, the spatial awareness concept presented in Chapter 3 is manifested in the advanced receptionist. In contrast to Figure 5.9, where modifications are described with regard to the system, Figure 5.10 depicts the relevance of the spatial awareness from a dialog point of view. The complete course of action as described in Chapter 3.2 is covered by the concept, which extends the robot's expressiveness during the whole encounter. With each behavior addressing one of the claims (cf. Chapter 3.3) in order to increase mutual understanding, Hypothesis H can be tested in comparison to the basic receptionist as a baseline scenario in the next chapter.

6. Impacts on a Robotic Encounter

In this chapter, the effectiveness of aforementioned spatial awareness strategies is examined. For this purpose, a user study is presented that evaluates the advanced strategies as introduced in Chapter 5 in contrast to the functionality of the basic receptionist as presented in Chapter 4. Such a comparison aspires to understand if modeling spatial properties based on humans has an advantage for the interaction and if the proposed implementation is feasible and thereby directly addresses Hypothesis H.

The next section subsequently presents the setup of the study including the physical arrangement of the robot, experimental conditions, questionnaire design, and an outline of participants. Afterwards, Section 6.2 statistically identifies differences in user ratings dependent on the specific behaviors the robot emits. Eventually in Section 6.3, observed effects are interpreted and related to the main hypothesis as well as individual claims established in Chapter 3.1 and 3.3.

6.1. Experimental Design

The immediate goal of the described experiment is to investigate aforementioned spatial strategies which not only consist of abstract guidelines but also recommendations for specific actions during an interaction. As a consequence, such behaviors have to be tested on a autonomously running live system and with naïve human partners. To inspect a complete encounter in a greater spatial domain, the interaction study has to be located in a room of appropriate size. As a second requirement, the questionnaire has to cover the complete encounter from the moment a participant enters the room on. As a direct consequence, participants must not come into contact with the robot prior to the study in order to prevent side effects of early appreciation or even familiarity.

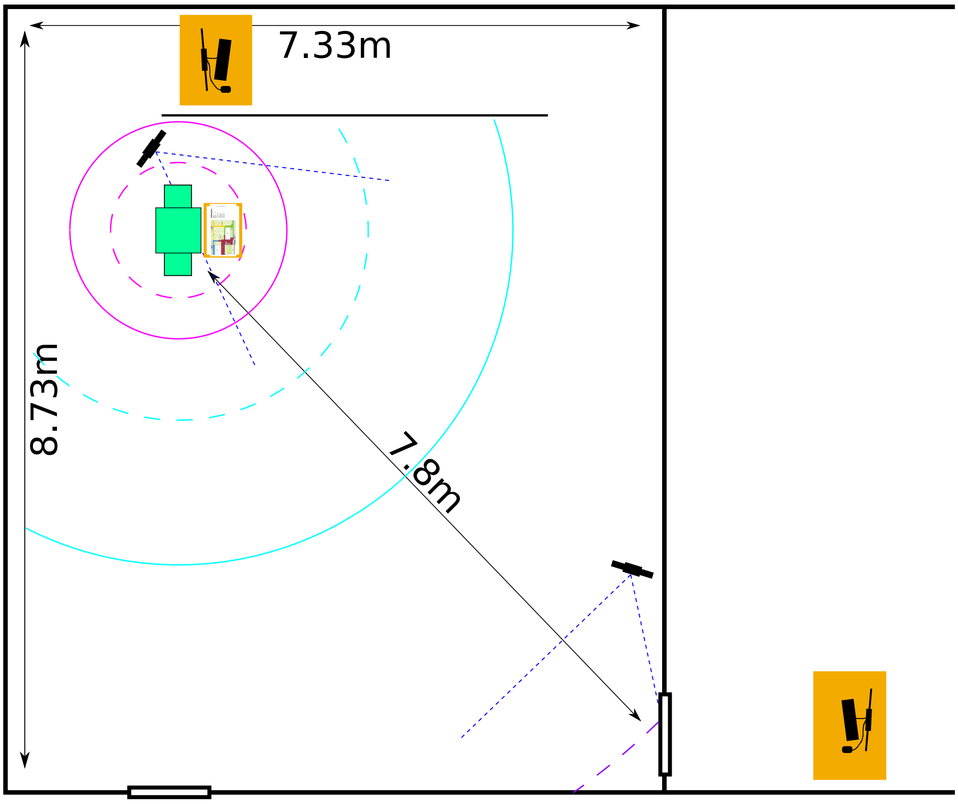


Figure 6.1.: Schematic depiction of the physical setup during the user study from above. The experiment room to the left contains the receptionist setup, two cameras, as well as the hidden investigator desk. In the second room, there is a workplace for participants of the study. Social distance classes (cf. Chap. 2.2.1) are given as bubbles surrounding the robot.

6.1.1. Physical Arrangement

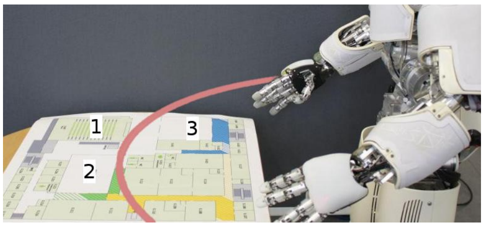
As an appropriate solution, the location for conducting the user study is set up as depicted in Figure 6.1. The experiment room to the left can be entered through an always open interconnecting door so that participant briefing can take place in the room to the right. Inside the room there is a large empty space in the center as well as the receptionist opposite to the designated entry point. Participants therefore have to approach the robot from afar and different strategies during pre-dialog phases can be explored. The social distance between robot and human at the entrance spans up until the beginnings of the far public class, so that all relevant segments have to be passed in order to engage in a face-to-face interaction. The robot is not yet oriented towards the door, so that it is forced to turn in order to redirect its attention towards a possible interaction partner. The receptionist is set up as pointed out in Chapter 4.2.2 with the floor plan directly in front of the robot on a small desk.

Behind a visual cover, there is a workspace for the investigator in close proximity to the robot where it is possible to observe the robot and shut it down immediately in emergency cases. One 3d-camera is employed at 175cm height behind the robot with the purpose to recognize humans while approaching as well as to register activities on the desk. Due to the limited range and resolution of the first camera, a secondary camera is used with regard to the initiation signal. It lies hidden on a desk between other hardware and acts as a detector for participants that enter the experiment room through the interconnecting door. Additionally, the experimenter takes notes on which types of gestures the robot carries out during the interaction in order to later comprehend the exact course of action and verify recorded log files.

6.1.2. Course of the Experiment

Participants are introduced to the scenario and their task at a computer inside the room to the right where they are yet unable to perceive the robot. During briefing they are instructed how to interact with the live robot system in terms of robot capabilities regarding speech and gesture. They are furthermore charged with a three-fold task they have to solve during the interaction. Namely, they have to ask for three different locations on the floor plan (cf. Appx. D) using different modalities in the following order:

1. Auditorium
2. Central inner courtyard
3. Inner courtyard at the edge



For retrieving the auditorium location, they should only use a verbal phrase, while users are requested to also incorporate pointing gestures towards the floor plan for both of the inner courtyards as to possibly generate obstructions during the interaction. One of the courtyards thereby is located at a much closer distance to the robot than the other, which lies at the outer limit of robot reach. For an enhanced reliability of speech recognition and a minimized participant distraction from their task, verbal input is not recorded via the wearable microphone but, with the help of utterance templates, typed in by the experimenter instead. After the interaction completes, they are asked to return to the briefing computer in order to fill out a questionnaire about the interaction. After clarifying the instructions, the investigator enters the experiment room in order to hide behind the visual cover, prepare the scenario and type in verbal utterances of the participants. Shortly afterwards, a signal is given to the participant who then follows into the room, approaches the robot, and solves the imposed tasks.

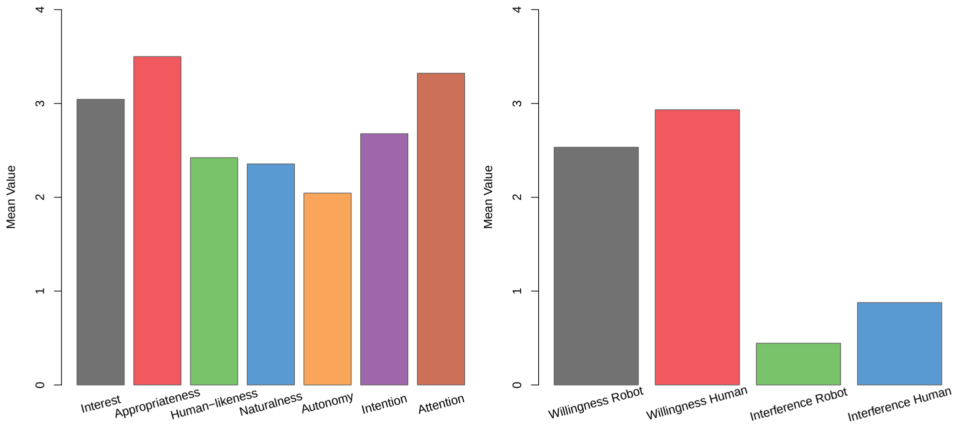
The questionnaire is available in German and English (cf. Appx. F) and starts with a self-assessment of robot and computer experience aimed to assure a simple entry into the questionnaire paired with a motivation to fill it out. Afterwards, questions regarding the opening of the interaction are asked, e.g. the robot's amount of willingness to interact during approaching and the moment when the participant thinks she has been first noticed by the robot the first time. On the next page, participants rate the robot behavior at dialog phase. Besides questions about the informative content of the different behavioral aspects, it is asked how much both agent's gestures interfere with each other and if the interaction changes between the second and third task, i.e. the two courtyards where the participant has to use gestures. At the end, personal information and feedback is retrieved in order estimate participant composition and attitude towards the experiment.

6.1.3. Participant Overview and Categorization

In total, 105 people have taken part in the study and received five Euros as a compensation for their efforts. 15 of them are completely excluded from the evaluation because they either have not followed the instructions or have experienced a faulty setup because of erroneous configurations or control by the operator. Of the remaining 90 participants 46 are female and 44 male. Participants' age ranges from 19 to 50 with a mean value of 26.1 years with an average self-assessed German knowledge of 3.99 on a (0-4) Likert scale and only a single rating below maximum. As the recruiting to a great part has taken place on campus, most participants are either students (78%) or scientific employees at the university (7%). Of the students, approximately 27% are enrolled in a subject related to natural or technical sciences. On a (0-4) scale, the participants' computer knowledge is solid with an average of 2.71, while in general they are relatively unfamiliar with robots other than those from movies as they rate their knowledge with 1.38 on average.

Two independent variables are setup as experimental conditions regarding interaction opening. At first, participants either experience a very distinct and therefore [S]trong initiation signals with the robot gaze directed straight towards their face when they enter the experiment room or a [W]eaker one, where the robot only turns 50% of the way before moving back to idling behavior. In both cases, however, only head and eye movements are employed. With the help of this modulation, it is aimed to draw conclusions about the style and impact of the initiation signal. As a second independent variable, participants during the study experience either, the approaching phase and opening signal are parametrized. One group therefore experiences only [R]andom movement during approaching as in the basic receptionist (Chapter 4.3). The integrated opening concept as described in Chapter 5.4 is exhibited in the [F]ull condition. In order to evaluate the importance of pro-active opening strategies, the same concept but with a [D]elayed robot greeting seven seconds after the participant enters the close social distance is demonstrated in a third group.

During dialog itself, another independent variable is employed in order to research the effect of *spatial prompting* strategies for conflict solving. In one group, the robot does not emit a pointing gesture at all towards the floor plan during explanation ([N]one group). Instead, it solely explains the way using real-world coordinates. All other groups experience robot gestures, while their obstruction strategies vary. In the [S]imple group, possible interferences cause an immediate cancellation of the gesture which reflects the basic receptionist. Spatial signals are only incorporated if the floor-plan is obstructed in interactions of the [P]rompting group. In parallel to this arrangement, participants are classified by interaction logs and annotations by the actual occurrence of a gesture being disturbed, resulting in groups describing if prompting is emitted, the gesture is discontinued, carried out regularly, or not even started by the robot.



(a) Ratings of robot properties

(b) Ratings of Willingness and Interference

Figure 6.2.: Participants' opinion about the interaction with the robot. In part (a) mean values of answers to the robot's properties are given. Part (b) gives the same numbers on the willingness as well as the amount of perceived interference by the counterpart.

6.2. Illustration of Results

This section illustrates the results obtained from the interaction study with the receptionist robot. At first, it provides a general overview of the obtained data from the questionnaire and impressions from individual trials. Then, more specific results are displayed and compared by experimental groups. Results regarding the integrated opening is presented first followed by a description of outcomes for the dialog phase. Interaction closing and departing is characterized with the help of experimenter observations and system logs in the end of this section.

In order to determine significant deviations between answers in the different conditions, a Kruskal-Wallis test (cf. Kruskal & Wallis 1952) is employed for ordinal data obtained from rating questions that have been answered on a (0-4) Likert scale. A χ^2 test of independence (cf. Pearson 1900) is utilized in case of nominal data (e.g. yes-no questions). Furthermore, a χ^2 goodness of fit test against equal probability is used to distinguish preference from chance for each decision inside each condition.

As an overview, Figure 6.2 gives an impression of ratings by evaluated participants (90) of the study. Every question regarding robot properties is rated above its arithmetic mean value by the participants as depicted in Figure 6.2a. The robot's interest in and its amount of attention towards the participant as well as the appropriateness of its behavior are rated best with a value of greater than three. The robot's autonomy is rated lowest with a mean value of 2.04. Part 6.2b displays the overall perceived willingness of the robot and oneself to engage in the interaction. With the own willingness rated higher, both values are above arithmetic mean with an average value 2.93 and 2.53. Interference instead is rated below one on average, while participants rate their own presence as more interfering (0.87) than the one of the robot (0.44). The comprehensive list of graphs containing study results as a whole and split by every independent variable can be found in Appendix G. In the following, only statistically significant differences between experiment conditions are pointed out. Excluded is a significant deviation in of age between the give up and prompting groups with averages of 27 and 24.94 years as it arguably does not imply any further consequences. Also, a difference in robot as well as computer experience between the groups [R] with averages of 1.54/2.86 and [F] with 0.9/2.35 and is left out. Such a difference may lead to an alteration but as the control group [R] rates itself as more experienced, at the most a decrease of effect size is expected.

6.2.1. Initiation Signal

The only significant difference between the weak and strong opening style can be found in answers to the question whether the robot greets the participant or not. Figure 6.3 accordingly displays participant answers to that question by initiation signal. While both groups do not significantly differ from overall answers to the question, participants experiencing a strong initiation signal negate the question more often as the other group. Furthermore, the decision for such an answer is reliably distinguishable from chance. This effect is supported by interaction logs that reveal a human-initiated greeting occurs in 93% of people in the [S] group as opposed to only 54% of people in the [W] group. In contrast to trial setups of the system, not a single participant shows behaviors of disorientation or searching the receptionist. Instead, everyone immediately turns towards the receptionist and approaches the desk.

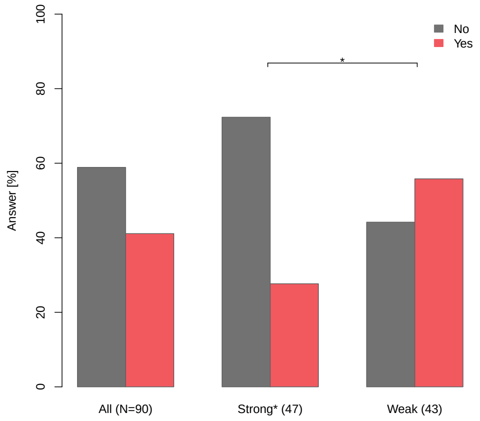
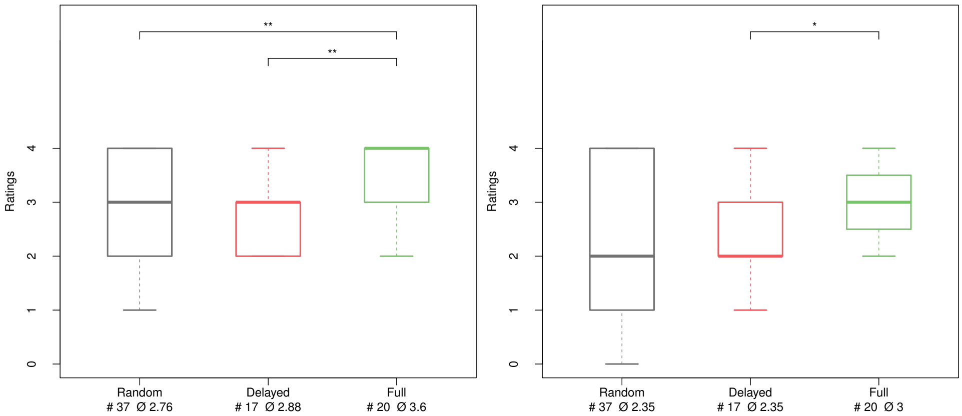


Figure 6.3.: Perceived robot greeting grouped by initiation signal (weak, strong). It is displayed if participants notice a greeting from the robot prior to their own utterance as answers to a yes/no question. Significance levels ($* := p < .05$) are given along the columns, as bars between the columns.

6.2.2. Integrated Interaction Opening

In this section, effects are described that occur between the different opening conditions (random, delayed, full). In none of the general questions regarding robot properties, any statistically relevant differences occur, so they are not displayed here. Instead, differences in the rated willingness to interact as well as the perceived means of opening appear, which are explained in the following. The first 26 participants of the study are left out from the results as they experienced a slightly different opening including the necessity to wear a microphone during approaching (also cf. Section 6.2.4).

At first, Figure 6.4 gives insights about participant ratings regarding their own willingness to take part in the interaction (Fig. 6.4a) as well the same perceived willingness of the robot (Fig. 6.4b). With an average of above three, participants rate their own as well as the robot's disposition highest in the full dynamic condition. Random only movements during approaching result in distinctly lower self-assessment of willingness compared to the dynamic condition. In the group experiencing dynamic attentive behavior with a delayed robot greeting, the willingness of both, robot and human is



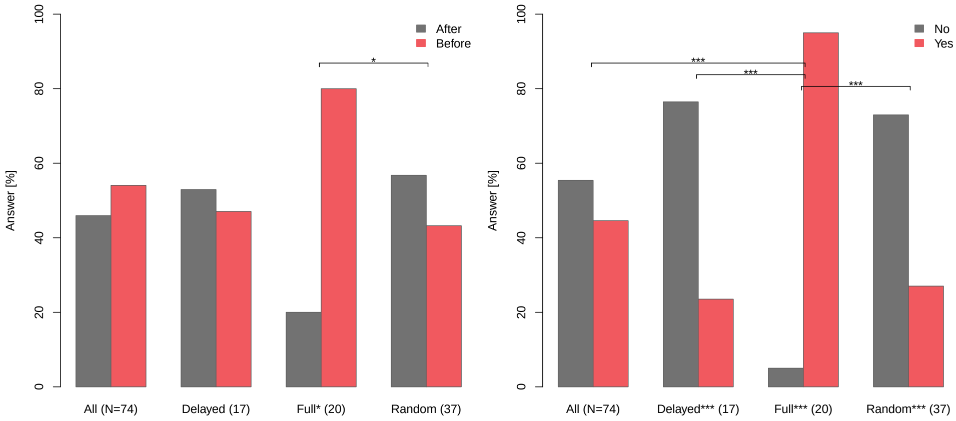
(a) Human willingness

(b) Robot willingness

Figure 6.4.: Box-plots containing ratings of willingness to take part in the interaction grouped by opening strategy (random, delayed, full). In part (a) opinions regarding the own willingness is displayed while part (b) gives the ratings of the perceived robot willingness. Significance levels ($* := p < .05$, $** := p < .01$) resulting from the independence test between the experimental conditions are given as bars between the boxes.

rated significantly lower as in the group receiving an immediate salutation upon entering the close social distance.

Regarding the perceived opening, the full dynamic group also differs from the others as depicted in Figure 6.5. At first, in this group, participants significantly earlier realize that the robot identifies them as a possible interaction partner compared to the random group. Group [D] and [R] on the other hand are not distinguishable from each other and the overall results (cf. Fig. 6.5a). Furthermore, a robot initiated greeting is noticed in the full dynamic group as opposed to all other groups. Instead, in delayed as well as in random-only conditions, a salutation by the robot is reliably not experienced as illustrated in Fig. 6.5b. This observation is reiterated in the interaction logs. Everyone in group [R] as well as 86% of group [D] in contrast to 28% in [F] take the initiative and greet the receptionist robot themselves.



(a) Perceived robot noticing*

(b) Perceived robot greet***

Figure 6.5.: Perceived robot awareness during approaching and dialog opening grouped by emitted behavior (random, delayed, full). In (a) it is displayed at which time during the encounter people think that the robot notices them. (b) reveals if participants notice a greeting initiated by the robot. Significance levels ($* := p < .05$, $** := p < .01$, $*** := p < .001$) are given along the columns, as bars between the columns, and at the title.

6.2.3. Face-to-Face Interaction

During the interaction with the robot at the table, in 56 trials, no interferences can be observed half of which no pointing gesture towards the map is carried out and the other half no hindrances occur. In total, 34 cases of disturbances of the robot's gesture in the shared *interaction space* occur. Four times the first gesture has been blocked, four participants cause alternatives to be triggered in both gestures, and 25 of them interfere the second pointing gesture. In 16 trials, the robot incorporates *prompting* strategies, while in the other 18, the gesture is aborted directly.

In the prompting case, which often involves multiple attempts, five times it is exhibited prior to the pointing gesture, while also five times, pointing has to be interrupted in order to allow for prompting. In two cases, the hand is switched for prompting and in two cases, the final pointing gesture is carried out with the left hand. Eight times, the gesture is aborted because

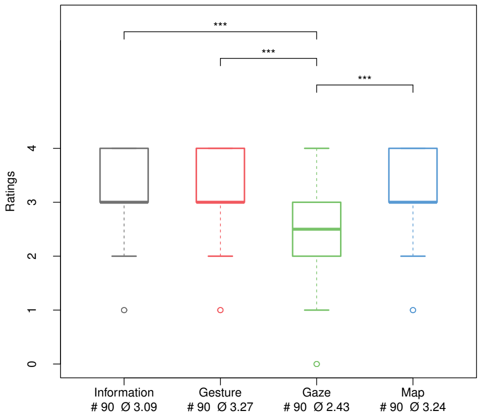


Figure 6.6.: Box-plot displaying participant ratings regarding the helpfulness of the receptionist explanations as a whole and grouped by information-bearing feature (gesture, gaze, map). Significance levels ($*** := p < .001$) resulting from the independence test between the features are given as bars in between the boxes.

the hindrance occurs at a close distance to the robot and one time, the gesture is aborted after prompting four times in a row.

Altogether, participants rate the helpfulness of information delivered by the receptionist with an average of 3.09 of 4. Figure 6.6 reveals that the robot gesture as well as the map with the floor plan on it strongly contribute to the informative content of the explanation. A significantly lower impact compared to the other components as well as in contrast to the overall content is given with the robot's gaze, with an average of 2.43.

Further answers to questions regarding the interaction at the table incorporating the map are in general displayed in two ways. One plot groups obtained results by the occurrence of obstruction, namely whether the robot gesture has been blocked, not been blocked, or otherwise no gesture has been carried out at all. A second plot emphasizes the incorporated robot strategy by further specifying the incorporated robot strategy. Besides the formerly employed categories of no block and no gesture, results are further split up whether the robot incorporates prompting as a strategy or simply gives up and discontinues the gesture.

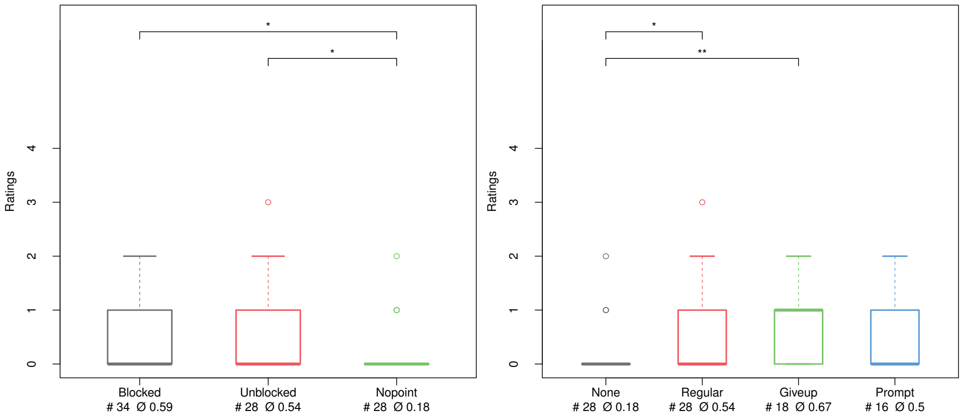
There are as well significant variations between the groups not only for the perceived interference caused by actions of the robot but also for the one participants emitted their own which is pointed out graphically in Figure 6.7. In the left part, i.e. 6.7a and 6.7b, a significant difference in robot as well as human interference can be observed between the conditions, in which the robot incorporates gestures towards the floor plan into its portfolio of actions whether they are blocked or not as opposed to directly gesturing towards the real world location (Nopoint). Additionally, only interference cause by the participant is significantly higher if the robot's gesture has been interrupted, while robot interference is not distinguishable in those cases.

In greater detail, Figure 6.7c and 6.7d reveal that as other than giving up and regular pointing, gestures involving spatial prompts increase the perceived interference produced by the robot to some degree but does not result in a significant change compared to no gesture. Participants on the other hand rate their own interference on robot gestures higher if the robot gestures on top of the floor plan with no observable difference between robot prompting and discontinuation.

Besides the perceived interference, Figure 6.8 demonstrates that participants also actively notice a change in the robot's way to carry out the pointing gesture towards the floor plan. In the case of no hindrances and no gesture, there is reliably claimed to be no change in the robot's behavior as pointed out in Figure 6.8a. In case of blocked gestures, participant answers differ significantly from the others as they detect a change more often. While spatial prompting as well as simple discontinuation can both be distinguished from the group with no gesture, the give-up strategy cannot reliably be separated from regular pointing in contrast to answers in the prompting group (cf. Fig. 6.8b).

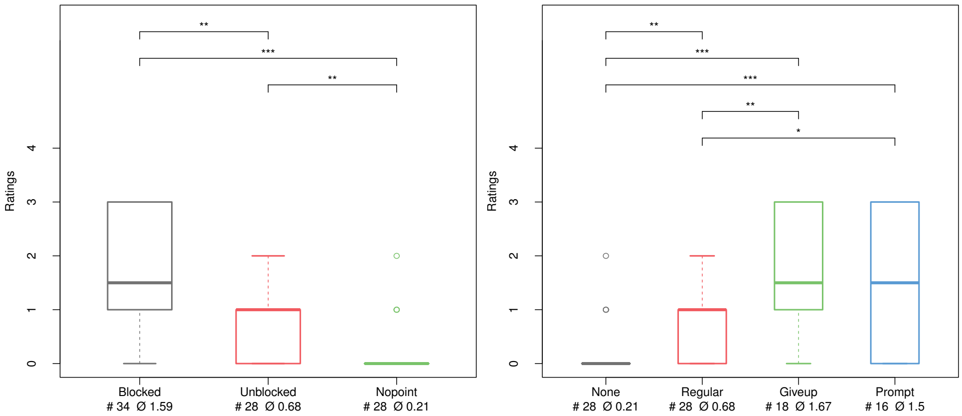
6.2.4. Ending the Encounter

During the interaction study, no direct measurements have been recorded regarding a closing signal or the departing phase. There are however observations of participant behavior that have a direct impact on the way of proceeding during the study. In the first 26 trials, the robot has not incorporated any active closing strategies. As a possible consequence, participants do not immediately return to the questionnaire after solving their assigned task. Instead, questions directed towards the experimenter whether the in-



(a) Robot-caused interference
By occurrence of obstruction

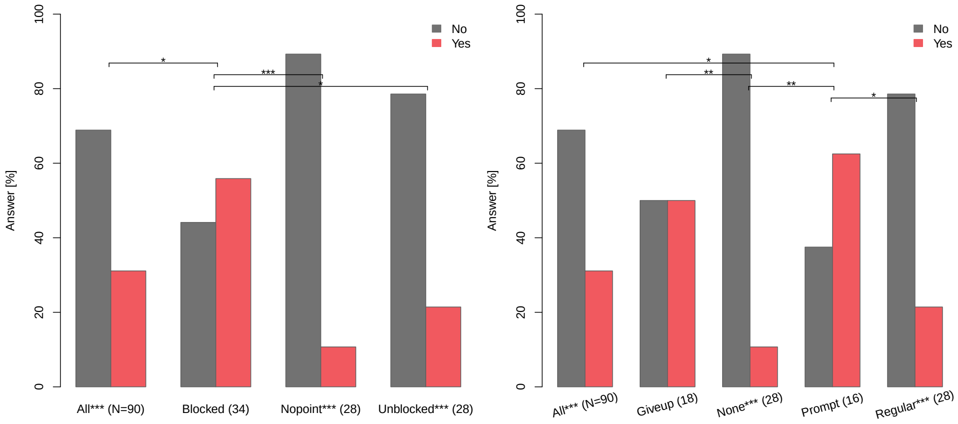
(b) Robot-caused interference
By experimental condition



(c) Human-caused interference
By occurrence of obstruction

(d) Human-caused interference
By experimental condition

Figure 6.7.: Box-plots containing ratings of perceived interference during the interaction. In part (a) and (b) interferences caused by the robot are displayed by the occurrence of obstructions on the left and experimental condition (regular, give up, prompt) on the right. (c) and (d) are grouped in the same way and give ratings of perceived interference that participants cause themselves. Significance levels ($* := p < .05$, $** := p < .01$, $*** := p < .001$) resulting from the independence test between the experimental conditions are given as bars between the boxes.



(a) By occurrence of obstruction***

(b) By experimental condition***

Figure 6.8.: Detected differences in the two consecutive explanations that employ a human pointing gesture. In (a) yes/no answers are grouped by the occurrence of an obstruction, while in (b) the same results are displayed depending on the experimental condition (regular, give up, prompt). Significance levels ($* := p < .05$, $** := p < .01$, $*** := p < .001$) are given along the columns, as bars between the columns, and at the title.

teraction is over amass at the end of the dialog. In cases where no such question arises, people stay in front of the robot after successful interaction until interrupted by the experimenter. As one possible solution, active closing as described in Chapter 5.6 has been added to the repertoire of robot capabilities leading to a successful disengagement behavior. As a further consequence, the first 26 people have been excluded in ratings regarding the robot's opening strategies.

6.3. Interpretation and Consequences

From the interactive study with the receptionist implications on Hypothesis H can be drawn as declared in Chapter 3.1. As a general feedback from the users, the system as a whole is rated excellently in terms of interest and attention. The amount of intention that the robot reveals as well as the overall appropriateness of its behaviors are also rated positive but to a slightly lesser extent. The system is apparently perceived as acting more or less human-like and natural, while the rating for autonomy only reaches mediocre levels. On the one hand, the latter might be influenced by the inevitable circumstance of knowing an operator in the experiment room albeit hidden behind a covering wall. On the other hand, users might not expect such a comprehensive set of behaviors from a receptionist robot. Only six participants however rate the robot behaviors as not autonomous at all.

Furthermore, it is noteworthy that ratings regarding robot properties do not differ significantly across all experimental conditions hinting at a self-explanatory and appropriate design of the interaction in addition to clear instructions during the study. Especially the way in which attention is expressed seems to work as intended. On the downside, possibly related to the overall high level of ratings or the relatively large number of intermixed experimental conditions, no clear distinction between behavioral categories can be drawn. Nevertheless, interesting differences emerge in questions especially addressing interaction instigation or conversation. In the following, these effects are interpreted and related to claims established in Chapter 3.3.

Initialization The experiment clearly shows that a short gaze as an initiation signal in still distant configurations works as intended and can lead to an immediate approaching behavior of a human towards the robot. The strength of such a gaze however has to be adjusted to the specific purpose of the individual setup. A stronger signal in the form of a direct gaze towards the head and eyes already causes human salutation utterances from a distance, whereas such an effect cannot be observed if the same gaze is exhibited in a less distinct fashion. If the robot is intended to demonstrate its potential during approaching including self initiated greetings, it is therefore inadvisable to use a very strong signal. In summary, Claim I can be approved based on the experiment with the advanced receptionist implemented as described in Chapter 5.1. An initiation gaze is recognized as an intentional communicative signal. Besides inducing an approaching behavior, it also often leads to salutation utterances from naïve users.

Approaching Furthermore, *proximity* and *orientation* dependent attentional behaviors are well suited as a robot strategy during the approaching phase. Results from the interaction experiment confirm the conclusions drawn from an earlier video study presented in Chapter 5.2. Although no differences in general ratings of robot properties can be determined, possibly caused by the large number of experimental conditions, opening specific questions allow for inference on the selected strategies. At first, participants are willing to take part in the interaction to a higher degree and also attest the robot the same. Secondly, participants are well aware that the robot notices them while they are approaching. As a result, Claim II can also be confirmed because the main purpose of leading the human into the interaction is fulfilled. Participants want to interact with the robot and notice that the robot is attentive and ready to be used.

Opening Answers to the questionnaire also indicate a tight coupling between attentive approaching and pro-active robot greeting as described in Chapter 5.3. Participants probably expect the robot to also open up the interaction if it raises its attention while they are coming closer. If that is not the case [D], no difference in the robot's and participant's willingness to interact in comparison to random-only movements [R] is identifiable. Above that, people only credit the robot the identification of themselves as an interaction partner during approaching if it does open up the interaction verbally. Possibly, participants interpret attentive behavior which besides in group [F] also is exhibited in [D] falsely as non-interactive because of an expected greeting which does not occur. Claim III therefore can be approved as it is not only appropriate to incorporate robot initiated opening but also required for a beneficial integration of attentive strategies prior to the dialog.

Dialog Pointing gesture and utterance both are major contributors to the amount of information delivered by the receptionist which hints at a well balanced interplay of gesture and speech. Gaze instead is meant as a supportive cue and is expected to have a lesser influence to that regard. The amount of overall obstructions that occur in the *interaction space* justifies the incorporation of behaviors on the robot targeted at circumventing or solving them conveniently. The integration of *spatial prompting* as described in Chapter 5.5 thereby qualifies as an appropriate strategy. Apparently, it does not effect ratings of robot properties in the conducted user study. Other than a great majority of blocks occurring in farther distance from the robot, no further influence of distance can be determined. Nonetheless, user experience in regard to perceived interference is altered. While movement alone results in some degree of interference, with prompting a relatively low amount of disturbances on the human, compared other robot movements seems to be caused. On the other hand, the participant is aware to cause collisions if the robot's gesture is aborted or if prompting is utilized, which confirms the effectiveness of prompting in terms of notifying of a desired pointing attempt and resolve ambiguities. Besides the an altered impression of interference, participants also are aware of the gesture being carried out differently if spatial prompting happens in contrast to cases where the robot simply gives up. As an authentic interpretation, prompting reveals aspects of the robot's inner state without causing major disturbances to the interaction and therefore successfully addresses Claim IV.

Closing and Departing As an implication from observations, an appropriate closing strategy seems to be an essential part of the interaction. An active strategy as outlined in Chapter 5.6 effectively closes the dialog. It seems to be equally expected as a proper greeting utterance causing the human to not really know how to proceed if not employed. As a result, Claim V can only be accepted provisorily due to the lack of questionnaire integration. In conjunction with proper non-verbal signals, disengagement behaviors of the human can be induced and accompanied. During the study, there has been no attempt of reengaging with the receptionist which makes it difficult to interpret the feasibility of decreasing attentiveness. However, it also hints to the correctness of Claim VI that nonverbal signals contribute largely to a human receding. The interconnection between verbal dialog opening and closing, as well as nonverbal signals furthermore supports the demand for an integrated and coordinated technique of emitting spatial signals with a social robot.

Conclusion In total, interesting conclusions can be drawn from the interaction study between naïve users and the receptionist. While overall differences between the experimental conditions have been expected to be more distinct, especially in property ratings, all claims can still be approved. A distance dependent attention strategy enables the robot to display its readiness plus its willingness to interact prior to a verbal conversation. By considering changes in spatial configurations it can be incorporated to effectively lead a visitor into and out of a conversation. Furthermore, a mutual dependency between the single behaviors has been demonstrated. During the dialog itself, spatial prompting supports the robot in actively expressing territorial needs without disturbing the human excessively.

As a conclusion, Hypothesis H can successfully be approved which implies that the incorporation of an integrated spatial awareness strategy on the receptionist leads to an enhancement in user experience. The study reveals that singals sent by the robot are in fact interpreted by participants as communicative acts that reveal information. As a consequence, the presented way of spatial awareness assists the human in developing more appropriate expectations about the situation by the emission of *spatial signals* in addition to its normal repertoire of actions. In conjunction with the construction of an own representation about shared spatial configurations, likewise with the analysis of signals, a contribution towards mutual understanding between human and robot is achieved.

6.4. Summary

In this chapter, the impact of spatial attentiveness on user experience has been evaluated. It describes a user study that installs the receptionist robot (cf. Chapter 4) in different experimental conditions in order to research the effect of social signals as described in Chapter 5 on the user. With the evaluation of this study reflecting on claims made in Chapter 3 which originate from research presented in Chapter 2, the main hypothesis can be confirmed in order to substantiate the approach that is taken in this work. This chapter hence integrates the previous chapters and thereby frames the overall contribution of this thesis.

For this purpose, Section 6.1 describes in detail how the study is configured and how the receptionist is set up in an appropriate location. Furthermore, it gives insights into the participant instructions and questionnaire that both are designed to enfold a two-part interaction consisting of approaching and departing as well as a face-to-face interaction at the receptionist desk. Section 6.2 afterwards elaborates on observations during the study and especially on significant effects that emerge from participants completing the questionnaire. The data is afterwards discussed in Section 6.3 resulting in conclusions that remain under the estimation but nonetheless allow for a positive assessment of stated claims and eventually the main hypothesis. As a result, implemented strategies provide methods for an unconscious extension of mutual mental models leading to a better understanding between human and robot.

With the evaluation of spatial awareness strategies accomplished, the overall contribution of thesis is reflected in the next chapter. Based on the presented work, a conclusion is drawn and opportunities for further enhancement of proposed behaviors and methods are presented.

7. Conclusion and Perspective

This chapter concludes the thesis with a reflection of the conducted work and its impact. Thereby, in Section 7.1 a brief recapitulation of the previous chapters is given with regard to their respective contributions. Afterwards, further insights that have been gained as a result of the chosen approach are presented in Section 7.2. Suggestions for follow-up research based on this thesis are pointed out in Section 7.3. Eventually, the thesis is closed in Section 7.4 with a short summary of the entire chapter.

7.1. Recapitulation of Contribution

As introduced in Chapter 1, robots are becoming more relevant as assistants in the daily life of humans. Consequently, humanoid robots in particular are more frequently equipped with social features to increase their usefulness in interactions with untrained users. A *social robot* thereby extends the communicative bandwidth with its human partner which commonly leads to an improved robustness of an interaction.

This thesis contributes to the field of social robotics with a novel approach towards human-like spatial awareness of a social robot. Contrary to existing attempts that typically focus on singular aspects, an entire encounter between one human and a robot is considered in order to guide a person through the complete interaction. Furthermore, the presented work is composed of a study-driven approach, in which theoretical concepts are at each stage of development related and adapted to *HRI*, implemented on an eligible system, and then verified with the help of user studies. In its general approach this thesis thus represents a desirable and at the same time viable methodology for developing a socially interactive robot.

The introduced concept is inspired and guided by the idea to facilitate a better understanding of the other by enriching each others representations (cf. Chap. 2). An extended exchange of nonverbal *spatial signals* thereby results in a more extensive *alignment* of the current situation between both agents. On the one hand, the interpretation of nonverbal signals allows the robot to raise its awareness of human intentions and thereby extends its repertoire of actions. On the other hand, raising the robot's nonverbal expressiveness due to the production of spatial signals allows humans to more efficiently assess the robot's capabilities and thereby automatically adapt their beliefs and expectations.

In order to cover the complete situation in an organized manner, key phases and signals are identified and addressed with a distinct strategy for the robot in Chapter 3. Furthermore, an integrated concept of spatial intelligence for a receptionist robot covering a holistic encounter with a naïve human interaction partner is presented. Although embedded in a specific scenario, such a concept is targeted at enhancing the general user experience with interactive robots. It is thus explicitly designed to foster mutual understanding with the help of communicative aspects of interpersonal configurations on different scales. Accordingly, each phase in such a human-robot encounter is investigated in a distinctive but at the same time interdependent way.

Aiming to investigate the effect of each strategy as well as the overall user experience, Chapter 4 presents a multi-modally interactive receptionist scenario bearing the possibility for enabling spatially aware functionality. Such a setup qualifies for researching spatial signals as it is sufficiently restricted but also adaptable enough to allow for experimental modifications in a controllable manner. After applying spatial awareness strategies to the receptionist robot and scenario in Chapter 5, a user study is conducted in order to identify differences in their ratings of the robot as well as their behavior in relation to a more basic interaction.

An interpretation of the evaluation in Chapter 6 indicates that the developed strategies have the potential to enhance an interaction with the robot especially by raising the user's willingness to participate in the scenario as well as outlining the perceived robot's awareness of the current situation. Furthermore, it is demonstrated that behavioral aspects are interconnected and cause expectations on the remaining parts of the encounter. As a result, the integrated concept of spatial awareness can be approved to fulfill its purpose and contribute to an enhanced mutual understanding. Apparently, social robots especially in real world applications can benefit substantially from utilizing the extended communicative capabilities that are enabled with the concept of spatial awareness.

7.2. Further Impact on Social Robotics

In addition to the contribution of a viable concept of spatial awareness for a social robot, further insights may be gained from this work. As a direct consequence of the proceeding, implications on the method of developing social robots emerge. At first, the need to extend behavioral modeling to cover the approaching and departing of an interaction partner is emphasized. A conversation itself is embedded in a more comprehensive encounter that has to be addressed and reflected when investigating the social capabilities of a robot. Not only in an appropriate communication range but also at further distances, robot signals, intended or not, shape the upcoming interaction.

Secondly, the study imposingly confirms an interdependency of behavioral aspects. Researchers have to estimate and evaluate possible effects of every robotic signal and their impact on subsequent parts of the interaction. With a study-driven approach, as presented in this work, changes of a robot's social capabilities can be investigated with respect to the current situation and later phases of the interaction. Video studies thereby allow for a more profound examination of singular aspects while at the same time detaching the user from the interaction, making real interactions equally important.

Thirdly, the presented concept greatly illustrates the opportunities of interpreting and emitting spatially relevant signals. Not only with mobile robots, such behaviors are highly viable for enhancing user experience on the one hand and interaction shaping possibilities for the researcher on the other hand. As each robot comprises of a spatial presence, explicitly conceptualizing an awareness considerably extends the robot's social capabilities. Besides equipping existing robots with such a concept, new robots should be constructed with regard to their spatial impression as well.

7.3. Subsequent Research on Spatial Awareness

Besides the implications for related research, the conducted research itself offers the potential for improvement and continuation. Although the study confirms positive effects of the integrated spatial awareness concept, it merely constitutes a starting point for more deeply examining each single phase of an encounter. Several possibilities for follow-up investigations are imaginable in order to verify and extend the conducted investigations.

For that purpose, it is suggested to refine the study with several smaller experiments focusing on particular aspects. For example, one could determine in detail, in which spatial configurations *prompting* behaviors should be enabled during a face-to-face interaction with a robot. In such an experiment that integrates results from Renner et al. (2014), the first independent variable could consist of the location on the map, i.e. close by the robot as opposed to close by the human and equally distant.

In addition, the robot's conflict resolving strategy with regard to its gesture can be investigated in greater detail. The only experimental distinction could consist of whether it incorporates a gesture or not and if it uses prompting in case of obstructions or not. This way, sharper focus can be laid on the dialog phase without having influences of any previously different interaction styles. Furthermore, a more structured investigation of active closing behaviors should be conducted in order to examine the exact timings which are most comfortable for the user. Following a similar mindset, more extensive insights could be gained on the conditions that influence dialog opening and *proxemics*. A study in the same scenario perhaps including robots of different sizes, cultural appearances, or personalities has the capability to gain additional knowledge in this regard.

7.4. Summary

To summarize, this chapter completes the thesis with a review that consolidates the theoretical background, derives concepts thereof, applies strategies onto an interactive scenario, and their evaluation and interpretation. It values the efforts of this thesis as a success and attributes it as a useful contribution to social robotics as it introduces a valid concept of spatial awareness enhancing human-robot interaction.

As insights for related research, it reasserts the importance of embedding human-robot interactions in a more comprehensive situation that range from an initial contact to the discontinuation of mutual awareness. It also confirms an interdependency of social signals and their generation of expectations from the human. Most importantly, it acknowledges the viability of spatial awareness in the field of social robotics. Finally, based on the synoptic approach of the study, possibilities for further research that focus even more on the details of specific phases which are delineated as a perspective.

Bibliography

Adams, B., Breazeal, C., Brooks, R. A., & Scassellati, B. (2000). Humanoid Robots: A New Kind of Tool. *Intelligent Systems and their Applications*, 15(4), 25–31, doi:10.1109/5254.867909. (cited on Pg. 3)

Albas, C. (1991). Proxemic Behavior: A Study of Extrusion. *The Journal of Social Psychology*, 131(5), 697–702, doi:10.1080/00224545.1991.9924653. (cited on Pg. 21)

Aleixo, A. C., Malhado, M., Ventura, R., & Lima, P. (2007). People Detection and Tracking in a Receptionist Robot. In *Conferência Portuguesa de Reconhecimento de Padrões*. (cited on Pg. 49, 50)

Allwood, J. (2002). Bodily Communication Dimensions of Expression and Content. In B. Granström, D. House, & I. Karlsson (Eds.), *Multimodality in Language and Speech Systems*, volume 19 of *Text, Speech and Language Technology* (pp. 7–26). Kluwer Academic Publishers, doi:10.1007/978-94-017-2367-1_2, isbn:9048160243. (cited on Pg. 13)

Argyle, M., Ed. (1969). *Social Interaction*. London: Transaction Publishers, isbn:0202309126. (cited on Pg. 3)

Argyle, M. & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 28(3), 289–304, doi:10.2307/2786027. (cited on Pg. 18)

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. `http://opencv.org`. (cited on Pg. 56)

Braningan, H. & Pearson, J. (2006). Alignment in Human-Computer Interaction. In K. Fischer (Ed.), *How People Talk to Computers, Robots, and Other Artificial Communication Partners* (pp. 140–156). (cited on Pg. 10)

Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, 42, 167–175, doi:10.1016/S0921-8890(02)00373-1. (cited on Pg. 14)

Breazeal, C. (2004). Function Meets Style: Insights From Emotion Theory Applied to HRI. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 34(2), 187–194, doi:10.1109/TSMCC.2004.826270. (cited on Pg. 15)

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of Nonverbal Communication on Efficiency and Robustness in Human-Robot Teamwork. In *International Conference on Intelligent Robots and Systems* (pp. 708–713). IEEE, doi:10.1109/IROS.2005.1545011. (cited on Pg. 2)

Breazeal, C., Takanishi, A., & Kobayashi, T. (2008). Social Robots that Interact with People. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook of Robotics* chapter 58, (pp. 1349–1369). Springer Berlin / Heidelberg, doi:10.1007/978-3-540-30301-5, isbn:9783540239574. (cited on Pg. 7)

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (2005). Real-time reference resolution by naive participants during a task-based unscripted conversation. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 153–171). MIT Press, isbn:0262701049. (cited on Pg. 14)

Brown-Schmidt, S. & Tanenhaus, M. K. (2004). Priming and alignment: Mechanism or consequence? *Behavioral and Brain Sciences*, 27(2), 193–194, doi:10.1017/S0140525X04250051. (cited on Pg. 9)

Brown-Schmidt, S. & Tanenhaus, M. K. (2008). Real-Time Investigation of Referential Domains in Unscripted Conversation: A Targeted Language Game Approach. *Cognitive Science*, 32(4), 643–684, doi:10.1080/03640210802066816. (cited on Pg. 61)

Brozzoli, C., Cardinali, L., Pavani, F., & Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia*, 48(3), 796–802, doi:10.1016/j.neuropsychologia.2009.10.009. (cited on Pg. 22)

Burghart, C., Holzapfel, H., Haeussling, R., & Breuer, S. (2007). Coding Interaction Patterns between Human and Receptionist Robot. In *International Conference on Humanoid Robots* (pp. 454–460). IEEE-RAS, doi:10.1109/ICHR.2007.4813909. (cited on Pg. 49, 50)

Cappelli, A., Giovannetti, E., & Federico, M. (2005). The Role of Communication in Human-Robot Interaction. In *RoboCare Workshop* (pp. 19–24). (cited on Pg. 15)

Caraian, S. & Kirchner, N. (2013). Influence of Robot-Issued Joint Attention Cues on Gaze and Preference. In *International Conference on Human-Robot Interaction* (pp. 95–96). ACM/IEEE, doi:10.1109/HRI.2013.6483518. (cited on Pg. 14)

Carbary, K. M., Frohning, E. E., & Tanenhaus, M. K. (2010). Context, Syntactic Priming, and Referential Form in an Interactive Dialogue Task: Implications for Models of Alignment. In *Conference of the Cognitive Science Society* (pp. 109–114). CSS. `https://mindmodeling.org/cogsci2010/papers/0017/`. (cited on Pg. 9)

Chinellato, E., Antonelli, M., Grzyb, B. J., & del Pobil, A. P. (2011). Implicit Sensorimotor Mapping of the Peripersonal Space by Gazing and Reaching. *Transactions on Autonomous Mental Development*, 3(1), 43–53, doi:10.1109/TAMD.2011.2106781. (cited on Pg. 23)

Ciolek, M. T. & Kendon, A. (1980). Environment and the Spatial Arrangement of Conversational Encounters. *Sociological Inquiry*, 50(3), 237–271, doi:10.1111/j.1475-682X.1980.tb00022.x. (cited on Pg. 21)

Ciolek, T. M. (1983). The proxemics lexicon: A first approximation. *Journal of Nonverbal Behavior*, 8(1), 55–79, doi:10.1007/BF00986330. (cited on Pg. 16)

Clark, H. H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39, doi:10.1016/0010-0277(86)90010-7. (cited on Pg. 8)

Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science*, 307(5712), 1082–5, doi:10.1126/science.1107799. (cited on Pg. 1)

Damm, O., Malchus, K., Hegel, F., Jaecks, P., Stenneken, P., Wrede, B., & Hielscher-fastabend, M. (2012). A computational model of emotional alignment. In *Workshop "Affective Computing" at German Conference on Artificial Intelligence*. `https://aiweb.techfak.uni-bielefeld.de/files/WS_EaC2011_odammEtAl.pdf`. (cited on Pg. 11)

Dautenhahn, K. (2007). Methodology and themes of human-robot interaction: a growing research field. *International Journal of Advanced Robotic Systems*, 4, 103–108, doi:10.5772/5702. (cited on Pg. 8)

Donath, J. (2011). Signals, cues and meaning. In *Signals, Truth and Design*. (unpublished manuscript). `http://judithdonath.net/judith/signalsTruthDesign.html`. (cited on Pg. 13)

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3-4), 177–190, doi:10.1016/S0921-8890(02)00374-3. (cited on Pg. 14)

Duffy, B. R. & Joue, G. (2004). I, Robot Being. In *Intelligent Autonomous Systems Conference*, doi:10.1.1.59.2833. (cited on Pg. 15)

Eimler, S. C., Krämer, N. C., & Pütten, A. M. V. D. (2010). Prerequisites for Human-Agent- and Human-Robot Interaction: Towards An Integrated Theory. In *European Meetings on Cybernetics and Systems Research*. (cited on Pg. 9)

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review*, 114(4), 864–886, doi:10.1037/0033-295X.114.4.864. (cited on Pg. 14)

Evers, V., et al. (2014). The Development and Real-World Deployment of FROG , the Fun Robotic Outdoor Guide. In *International Conference on Human-Robot Interaction* (pp. 100). ACM/IEEE, doi:10.1145/2559636.2559649. (cited on Pg. 47)

Eyssel, F., Hegel, F., Horstmann, G., & Wagner, C. (2010). Anthropomorphic inferences from emotional nonverbal cues: A case study. In *International Symposium on Robot and Human Interactive Communication* (pp. 646–651). IEEE, doi:10.1109/ROMAN.2010.5598687. (cited on Pg. 14)

Fink, G. A. (1999). Developing HMM-Based Recognizers with ESMERALDA. In V. Matousek, P. Mautner, J. Ocelíková, & P. Sojka (Eds.), *Text, Speech and Dialogue*, volume 1692 of *Lecture Notes in Computer Science* (pp. 229–234). Springer Berlin Heidelberg, doi:10.1007/3-540-48239-3_42, isbn:978-3-540-66494-9. (cited on Pg. 56)

Fischer, K. (2011). How People Talk with Robots: Reduce User Uncertainty. *AI Magazine*, 32(4), 31–38, doi:10.1609/aimag.v32i4.2377. (cited on Pg. 15)

Fischer, K., Lohan, K., Rohlfing, K., & Foth, K. (2014). Partner Orientation in Asymmetric Communication: Evidence from contingent robot response. In *Workshop on Humans and Robots in Asymmetric Interactions at the International Conference on Human-Robot Interaction*. `https://aiweb.techfak.uni-bielefeld.de/hri2014_workshop_asymmetric_interactions/FinalSubmissions/Fischer.pdf`. (cited on Pg. 8)

Fitzpatrick, P., Metta, G., & Natale, L. (2008). Towards long-lived robot genes. *Robotics and Autonomous Systems*, 56(1), 29–45, doi:10.1016/j.robot.2007.09.014. (cited on Pg. 57)

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003a). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3), 143–166, doi:10.1016/S0921-8890(02)00372-X. (cited on Pg. 4, 28)

Fong, T., Thorpe, C., & Baur, C. (2003b). Collaboration, Dialogue, and Human-Robot Interaction. In R. A. Jarvis & A. Zelinsky (Eds.), *Robotics Research*, volume 6 of *Springer Tracts in Advanced Robotics* (pp. 255–266). Springer Berlin / Heidelberg, doi:10.1007/3-540-36460-9_17, isbn:978-3-540-00550-6. (cited on Pg. 13)

Gan, T., Wong, Y., Zhang, D., & Kankanhalli, M. S. (2013). Temporal Encoded F-formation System for Social Interaction Detection. In *International Conference on Multimedia* (pp. 937–946). ACM, doi:10.1145/2502081.2502096. (cited on Pg. 21)

Ghosh, M. & Kuzuoka, H. (2013). A Trial Attempt by a Museum Guide Robot to Engage and Disengage the Audience on Time. In *Winter International Conference on Engineering and Technology* (pp. 18–22). Atlantic Press, doi:10.2991/wiet-13.2013.5. (cited on Pg. 32)

Ghosh, M. & Kuzuoka, H. (2014). An Ethnomethodological Study of a Museum Guide Robot's Attempt at Engagement and Disengagement. *Journal of Robotics*, doi:10.1155/2014/876439. (cited on Pg. 88)

Gockley, R., et al. (2005). Designing Robots for Long-Term Social Interaction. In *International Conference on Intelligent Robots and Systems* (pp. 1338–1343). IEEE, doi:10.1109/IROS.2005.1545303. (cited on Pg. 49, 50)

Goerick, C., Wersing, H., Mikhailova, I., & Dunn, M. (2005). Peripersonal space and object recognition for humanoids. In *International Conference on Humanoid Robots* (pp. 387–392). IEEE-RAS, doi:10.1109/ICHR.2005.1573598. (cited on Pg. 23)

Goodrich, M. A. & Schultz, A. C. (2007). Human-Robot Interaction: A Survey. *Foundations and Trends in Human-Computer Interaction*, 1(3), 203–275, doi:10.1561/1100000005. (cited on Pg. 4, 13)

Green, A. & Hüttenrauch, H. (2006). Making a Case for Spatial Prompting in Human-Robot Communication. In *International Conference on Language Resources and Evaluation Workshop: Multimodal Corpora: From ultimodal behaviour theories to usable models*. `ftp://ftp.nada.kth.se/IPLab/TechReports/HCI-24.pdf`. (cited on Pg. 24, 31)

Greenberg, S., Hornbæ k, K., Quigley, A., Reiterer, H., & Rädle, R. (2014). Proxemics in Human-Computer Interaction. *Dagstuhl Reports*, 3(11), 29–57, doi:10.4230/DagRep.3.11.29. (cited on Pg. 20)

Gross, H.-M., et al. (2009). TOOMAS: Interactive Shopping Guide Robots in Everyday Use - Final Implementation and Experiences from Long-term Field Trials. In *International Conference on Intelligent Robots and Systems* (pp. 2005–2012). IEEE/RSJ, doi:10.1109/IROS.2009.5354497. (cited on Pg. 1)

Guitton, D. & Volle, M. (1987). Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *Journal of neurophysiology*, 58, 427–459. `http://ncbi.nlm.nih.gov/pubmed/3655876`. (cited on Pg. 83)

Hackel, M., Schwope, S., Fritsch, J., Wrede, B., & Sagerer, G. (2005). A Humanoid Robot Platform Suitable for Studying Embodied Interaction. In *International Conference on Intelligent Robots and Systems* (pp. 56–61). IEEE/RSJ, doi:10.1109/IROS.2005.1544959. (cited on Pg. 66)

Hall, E. T. (1963). A System for the Notation of Proxemic Behavior. *American Anthropologist*, 65(5), 1003–1026, doi:10.1525/aa.1963.65.5.02a00020. (cited on Pg. 17)

Hall, E. T. (1966). *The Hidden Dimension*. Garden City: Doubleday, isbn:0385084765. (cited on Pg. 16, 17, 81, 82)

Hall, E. T., et al. (1968). Proxemics. *Current Anthropology*, 9(2/3), 83, doi:10.1086/200975. (cited on Pg. 16, 17, 19, 28, 30)

Hashimoto, T., Hiramatsu, S., Tsuji, T., & Kobayashi, H. (2007). Realization and Evaluation of Realistic Nod with Receptionist Robot SAYA. In *International Symposium on Robot and Human Interactive Communication* (pp. 326–331). IEEE, doi:10.1109/ROMAN.2007.4415103. (cited on Pg. 49, 50)

Heenan, B., Greenberg, S., Aghelmanesh, S., & Sharlin, E. (2013). *Designing Social Greetings and Proxemics in Human Robot Interaction*. Technical report, Department of Computer Science, University of Calgary. `http://grouplab.cpsc.ucalgary.ca/grouplab/uploads/Publications/Publications/2013-HRIGreetings.Report2013-1044-11.pdf`. (cited on Pg. 22)

Hegel, F., Eyssel, F., & Wrede, B. (2010). The Social Robot 'Flobi': Key Concepts of Industrial Design. In *International Symposium on Robot and Human Interactive Communication* (pp. 107–112). IEEE, doi:10.1109/ROMAN.2010.5598691. (cited on Pg. 66)

Hegel, F., Gieselmann, S., Peters, A., Holthaus, P., & Wrede, B. (2011). Towards a Typology of Meaningful Signals and Cues in Social Robotics. In *International Symposium on Robot and Human Interactive Communication* (pp. 72–78). IEEE, doi:10.1109/ROMAN.2011.6005246. (cited on Pg. 13)

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M., & Sagerer, G. (2009). Understanding Social Robots. In *International Conferences on Advances in Computer-Human Interactions* (pp. 169–174). IEEE, doi:10.1109/ACHI.2009.51. (cited on Pg. 3)

Hoffman, G. & Breazeal, C. (2007). Effects of Anticipatory Action on Human-Robot Teamwork Efficiency, Fluency, and Perception of Team. In *International Conference on Human-Robot Interaction* (pp. 1–8). ACM/IEEE, doi:10.1145/1228716.1228718. (cited on Pg. 15)

Holmes, N. P., Calvert, G. A., & Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience Letters*, 372(1-2), 62–67, doi:10.1016/j.neulet.2004.09.024. (cited on Pg. 22)

Holmes, N. P. & Spence, C. (2004). The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, 5(2), 94–105, doi:10.1007/s10339-004-0013-3. (cited on Pg. 22)

Holthaus, P., Pitsch, K., & Wachsmuth, S. (2011). How Can I Help? - Spatial Attention Strategies for a Receptionist Robot. *International Journal of Social Robotics*, 3(4), 383–393, doi:10.1007/s12369-011-0108-9. (cited on Pg. 65)

Holthaus, P. & Wachsmuth, S. (2012). Active Peripersonal Space for More Intuitive HRI. In *International Conference on Humanoid Robots* (pp. 508–513). IEEE-RAS, doi:10.1109/HUMANOIDS.2012.6651567. (cited on Pg. 33)

Holthaus, P. & Wachsmuth, S. (2014). The Receptionist Robot. In *International Conference on Human-Robot Interaction* (pp. 329–329). ACM/IEEE, doi:10.1145/2559636.2559784. (cited on Pg. 49)

Holzapfel, H. & Waibel, A. (2007). Behavior Models for Learning and Receptionist Dialogs. In *Conference of the International Speech Communication Association* (pp. 2189–2192). `http://www.isca-speech.org/archive/archive_papers/interspeech_2007/i07_2189.pdf`. (cited on Pg. 49, 50)

Hüttenrauch, H., Eklundh, K., & Green, A. (2006). What's in the gap? Interaction transitions that make HRI work. In *International Symposium on Robot and Human Interactive Communication* (pp. 123–128). IEEE, doi:10.1109/ROMAN.2006.314405. (cited on Pg. 24)

Jung, J., Kanda, T., & Kim, M.-S. (2013). Guidelines for Contextual Motion Design of a Humanoid Robot. *International Journal of Social Robotics*, 5(2), 153–169, doi:10.1007/s12369-012-0175-6. (cited on Pg. 15, 31)

Kahn, P. H., Gill, B. T., Reichert, A. L., Kanda, T., Ishiguro, H., & Ruckert, J. H. (2010). Validating interaction patterns in HRI. In *International Conference on Human-Robot Interaction* (pp. 183–184). ACM/IEEE, doi:10.1109/HRI.2010.5453205. (cited on Pg. 8)

Kampe, K. K. W., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Psychology: Reward value of attractiveness and gaze. *Nature*, 413(589), doi:10.1038/35098149. (cited on Pg. 65)

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, doi:10.1207/s15327051hci1901&2_4. (cited on Pg. 3)

Kaneko, K., Kanehiro, F., Morisawa, M., Akachi, K., Miyamori, G., Hayashi, A., & Kanehira, N. (2011). Humanoid robot HRP-4 - Humanoid robotics platform with lightweight and slim body. In *International Conference on Intelligent Robots and Systems* (pp. 4400–4407). IEEE/RSJ, doi:10.1109/IROS.2011.6094465. (cited on Pg. 1)

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22–63, doi:10.1016/0001-6918(67)90005-4. (cited on Pg. 69)

Kendon, A. (1990). *Conducting interaction: Patterns of social behavior in focused encounters*. Cambridge University Press, isbn:0521389380. (cited on Pg. 21, 23, 29, 30, 64, 65, 79, 80)

Kendon, A. (1997). Gesture. *Annual Review of Anthropology*, 26, pp. 109–128. `http://www.jstor.org/stable/2952517`. (cited on Pg. 8, 31)

Kirby, R., Simmons, R., & Forlizzi, J. (2009). COMPANION: A Constraint Optimizing Method for Person-Acceptable Navigation. In *International Symposium on Robot and Human Interactive Communication* (pp. 607–612). IEEE, doi:10.1109/ROMAN.2009.5326271. (cited on Pg. 20)

Knapp, M., Hall, J., & Horgan, T. (2013). *Nonverbal communication in human interaction*. Cengage Learning, isbn:1133311598. (cited on Pg. 11)

Koay, K. L., Syrdal, D. S., Walters, M. L., & Dautenhahn, K. (2007). Living with Robots: Investigating the Habituation Effect in Participants: Preferences During a Longitudinal Human-Robot Interaction Study. In *International Symposium on Robot and Human Interactive Communication* (pp. 564–569). IEEE, doi:10.1109/ROMAN.2007.4415149. (cited on Pg. 20)

Kopp, S. & Wachsmuth, I. (2004). Synthesizing multimodal utterances for conversational agents. *Computer Animation and Virtual Worlds*, 15(1), 39–52, doi:10.1002/cav.6. (cited on Pg. 8)

Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE*, 3(7), doi:10.1371/journal.pone.0002597. (cited on Pg. 3)

Krämer, N. C. (2005). Social Communicative Effects of a Virtual Program Guide. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, & T. Rist (Eds.), *Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science* (pp. 442–453). Springer Berlin Heidelberg, doi:10.1007/11550617_37, isbn:978-3-540-28738-4. (cited on Pg. 32)

Krauss, R. M. & Pardo, J. S. (2004). Is alignment always the result of automatic priming? *Behavioral and Brain Sciences*, 27(2), 203–204, doi:10.1017/S0140525X0436005X. (cited on Pg. 9)

Kruskal, W. H. & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583–621. `http://www.jstor.org/stable/2280779`. (cited on Pg. 98)

Kuzuoka, H., Suzuki, Y., Yamashita, J., & Yamazaki, K. (2010). Reconfiguring spatial formation arrangement by robot body orientation. In *International Conference on Human-Robot Interaction* (pp. 285–292). ACM/IEEE, doi:10.1145/1734454.1734557. (cited on Pg. 21)

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human-Robot Interaction. *Journal* *of Communication*, 56, 754–772, doi:10.1111/j.1460-2466.2006.00318.x. (cited on Pg. 3)

Lee, M. K., Kiesler, S., & Forlizzi, J. (2010). Receptionist or Information Kiosk: How Do People Talk With a Robot? In *Computer Supported Cooperative Work* (pp. 31–40). ACM, doi:10.1145/1718918.1718927. (cited on Pg. 3, 49, 50)

Leite, I., Martinho, C., & Paiva, A. (2013). Social Robots for Long-Term Interaction: A Survey. *International Journal of Social Robotics*, 5(2), 291–308, doi:10.1007/s12369-013-0178-y. (cited on Pg. 3)

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 140, 1–55. `http://psycnet.apa.org/psycinfo/1933-01885-001`. (cited on Pg. 36, 72, 97, 98)

Lloyd, D. M. (2009). The space between us: a neurophilosophical framework for the investigation of human interpersonal space. *Neuroscience & Biobehavioral Reviews*, 33(3), 297–304. `http://www.ncbi.nlm.nih.gov/pubmed/18926850`. (cited on Pg. 22, 33)

Lohse, M. (2011). The role of expectations and situations in human-robot interaction. In K. Dautenhahn & J. Saunders (Eds.), *New Frontiers in Human-Robot Interaction*, volume 2 of *Advances in Interaction Studies* (pp. 35–56). John Benjamins, isbn:9789027204554. (cited on Pg. 14)

Lohse, M., Hegel, F., Swadzba, A., Rohlfing, K., Wachsmuth, S., & Wrede, B. (2007). What can I do for you? Appearance and Application of Robots. In *Workshop "The Reign of Catz and Dogz? The role of virtual creatures in a computerised society" at Artificial Intelligence and Simulation of Behaviour Convention*. `https://aiweb.techfak.uni-bielefeld.de/files/papers/Lohse2007-AAR.pdf`. (cited on Pg. 14, 47)

Lömker, F., Wrede, S., Hanheide, M., & Fritsch, J. (2006). Building Modular Vision Systems with a Graphical Plugin Environment. In *International Conference on Vision Systems* (pp. 2–2). IEEE, doi:10.1109/ICVS.2006.18. (cited on Pg. 55)

Lütkebohle, I., Hegel, F., Schulz, S., Hackel, M., Wrede, B., Wachsmuth, S., & Sagerer, G. (2010). The Bielefeld Anthropomorphic Robot

Head "Flobi". In *International Conference on Robotics and Automation* (pp. 3384–3391). IEEE-RAS, doi:10.1109/ROBOT.2010.5509173. (cited on Pg. 66)

Lütkebohle, I., Peltason, J., Schillingmann, L., Elbrechter, C., Wrede, B., Wachsmuth, S., & Haschke, R. (2009a). The curious robot - structuring interactive robot learning. In *International Conference on Robotics and Automation* (pp. 4156–4162). IEEE, doi:10.1109/ROBOT.2009.5152521. (cited on Pg. 4, 22, 31)

Lütkebohle, I., Philippsen, R., Pradeep, V., Marder-Eppstein, E., & Wachsmuth, S. (2011). Generic middleware support for coordinating robot software components: The Task-State-Pattern. *Journal of Software Engineering for Robotics*, 2(1), 20–39. `http://joser.unibg.it/index.php?journal=joser&page=article&op=view&path=41`. (cited on Pg. 55)

Lütkebohle, I., Schaefer, J., & Wrede, S. (2009b). Facilitating Re-Use by Design: A Filtering, Transformation, and Selection Architecture for Robotic Software Systems. In *Workshop on "Software Engineering for Robotics" at the International Conference on Robots and Automation*. `https://aiweb.techfak.uni-bielefeld.de/files/sdir09.pdf`. (cited on Pg. 57)

Makatchev, M., Simmons, R., Sakr, M., & Ziadee, M. (2013). Expressing Ethnicity through Behaviors of a Robot Character. In *International Conference on Human-Robot Interaction* (pp. 357–364). ACM/IEEE. `http://dl.acm.org/citation.cfm?id=2447681`. (cited on Pg. 49, 51)

Marshall, P., Rogers, Y., & Pantidi, N. (2011). Using F-formations to analyse spatial patterns of interaction in physical environments. In *Conference on Computer Supported Cooperative Work* (pp. 445–454). ACM Press, doi:10.1145/1958824.1958893. (cited on Pg. 21)

Mavridis, N. (2014). A Review of Verbal and Non-Verbal Human-Robot Interactive Communication. *arXiv:1401.4994 [cs.RO]*. `http://arxiv.org/pdf/1401.4994v1`. (cited on Pg. 8)

Maynard-Smith, J. & Harper, D. (2004). *Animal signals*. Oxford University Press, isbn:0198526857. (cited on Pg. 13)

Menenti, L., Pickering, M. J., & Garrod, S. C. (2012). Toward a neural basis of interactive alignment in conversation. *Frontiers in Human Neuroscience*, 6(185), doi:10.3389/fnhum.2012.00185. (cited on Pg. 9)

Metta, G., et al. (2010). The iCub humanoid robot: an open-systems platform for research in cognitive development. *Neural Networks*, 23(8-9), 1125–1134, doi:10.1016/j.neunet.2010.08.010. (cited on Pg. 1, 34, 53)

Michalowski, M., Sabanovic, S., & Simmons, R. (2006). A spatial model of engagement for a social robot. In *International Workshop on Advanced Motion Control* (pp. 762–767). IEEE, doi:10.1109/AMC.2006.1631755. (cited on Pg. 49, 50)

Mondada, L. (2009). 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, doi:10.1016/j.pragma.2008.09.019. (cited on Pg. 65)

Moore, C. & Dunham, P., Eds. (1995). *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum Associates Inc., isbn:0-8058-1437-X. (cited on Pg. 14)

Mundy, P. & Newell, L. (2007). Attention, Joint Attention, and Social Cognition. *Current directions in psychological science*, 16(5), 269–274, doi:10.1111/j.1467-8721.2007.00518.x. (cited on Pg. 14)

Nguyen, N. & Wachsmuth, I. (2008). Towards a computational model of an agent's awareness of body and space. In *Symposium "Emotional Computer Systems and Interfaces" at International Conference on Cognitive Science* (pp. 553–554). `http://www.techfak.uni-bielefeld.de/~nnguyen/papers/nnguyen_iwachsmuth_2008.pdf`. (cited on Pg. 23)

Nguyen, N. & Wachsmuth, I. (2011). From body space to interaction space-modeling spatial cooperation for virtual humans. In *International Conference on Autonomous Agents and Multiagent Systems* (pp. 1047–1054). `http://ifaamas.org/Proceedings/aamas2011/papers/D7_B79.pdf`. (cited on Pg. 23)

Nisimura, R., Uchida, T., Lee, A., Saruwatari, H., Shikano, K., & Matsumoto, Y. (2002). ASKA: Receptionist Robot with Speech Dialogue System. In *International Conference on Intelligent Robots and Systems* (pp. 1314–1319). IEEE/RSJ, doi:10.1109/IRDS.2002.1043936. (cited on Pg. 49)

Nomikou, I., Pitsch, K., & Rohlfing, K. J. (2013). Asymmetry and adaptation in social interaction: A micro-analytic perspective. *Interaction Studies*, 14(2), 7–12, doi:10.1075/is.14.2.001nom. (cited on Pg. 7)

Pacchierotti, E., Christensen, H. I., & Jensfelt, P. (2006). Evaluation of passing distance for social robots. In *International Symposium on Robot and Human Interactive Communication* (pp. 315–320). IEEE, doi:10.1109/ROMAN.2006.314436. (cited on Pg. 20)

Paepcke, S. & Takayama, L. (2010). Judging a Bot By Its Cover: An Experiment on Expectation Setting for Personal Robots. In *International Conference on Human-Robot Interaction* (pp. 45–52). ACM/IEEE, doi:10.1109/HRI.2010.5453268. (cited on Pg. 15)

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382–2393, doi:10.1121/1.2178720. (cited on Pg. 9)

Pattacini, U. (2011). *Modular Cartesian Controllers for Humanoid Robots: Design and Implementation on the iCub.* PhD thesis, University of Genoa. `http://www.iit.it/images/images/icub-facility/docs/theses/pattacini.pdf`. (cited on Pg. 57, 83)

Pattacini, U., Nori, F., Natale, L., Metta, G., & Sandini, G. (2010). An Experimental Evaluation of a Novel Minimum-Jerk Cartesian Controller for Humanoid Robots. In *International Conference on Intelligent Robots and Systems* (pp. 1668–1674). IEEE/RSJ, doi:10.1109/IROS.2010.5650851. (cited on Pg. 57)

Patterson, M. L. (1982). A Sequential Functional Model of Nonverbal Exchange. *Psychological Review*, 89(3), 231–249, doi:10.1037//0033-295X.89.3.231. (cited on Pg. 11, 64)

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, 50(302), 157–175, doi:10.1080/14786440009463897. (cited on Pg. 37, 98)

Peltason, J., Lütkebohle, I., Wrede, B., & Hanheide, M. (2009). Mixed-Initiative in Interactive Robotic Learning. In *Workshop on "Improving Human-Robot Communication with Mixed-Initiative and Context-Awareness" at International Symposium on Robot and Human Interactive Communication*. `http://ceur-ws.org/Vol-693/paper5.pdf`. (cited on Pg. 32)

Peltason, J., Rieser, H., Wachsmuth, S., & Wrede, B. (2013). On Grounding Natural Kind Terms in Human-Robot Communication. *Künstliche Intelligenz*, 27(2), 107–118, doi:10.1007/s13218-013-0240-9. (cited on Pg. 9)

Peltason, J. & Wrede, B. (2010). Pamini: A framework for assembling mixed-initiative human-robot interaction from generic interaction patterns. In *Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 229–232). `http://dl.acm.org/citation.cfm?id=1944546`. (cited on Pg. 56)

Perelle, I. B. & Ehrman, L. (1994). An International Study of Human Handedness: The Data. *Behavior Genetics*, 24(3), 217–227. `http://www.ncbi.nlm.nih.gov/pubmed/7945152`. (cited on Pg. 34)

Peters, A. (2011). Small movements as communicational cues. In *"Pioneers" Workshop at International Conference on Human-Robot Interaction*. `https://aiweb.techfak.uni-bielefeld.de/files/apeters_HRI_Pioneers.pdf`. (cited on Pg. 24)

Peters, A. (2012). *Spatial Coordination - Human and Robotic Communicative Whole-Body Motions in Narrow Passages*. PhD thesis, Bielefeld University. `http://pub.uni-bielefeld.de/publication/2594360`. (cited on Pg. 24)

Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, 27(2), 169–190; discussion 190–226, doi:10.1017/S0140525X04000056. (cited on Pg. 9, 10)

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 *Robot and Human Interactive Communication* (pp. 985–991). IEEE, doi:10.1109/ROMAN.2009.5326167. (cited on Pg. 22)

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, doi:10.1075/is.14.2.06pit. (cited on Pg. 13)

Pitsch, K., Wrede, S., Seele, J.-c., & Süssenbach, L. (2011). Attitude of german museum visitors towards an interactive art guide robot. In *International Conference on Human-Robot Interaction* (pp. 227–228). ACM/IEEE, doi:10.1145/1957656.1957744. (cited on Pg. 4)

Remland, M. S., Jones, T. S., & Brinkman, H. (1995). Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. *The Journal of Social Psychology*, 135(3), 281–297, doi:10.1080/00224545.1995.9713958. (cited on Pg. 17)

Renner, P., Pfeiffer, T., & Wachsmuth, I. (2014). Spatial References with Gaze and Pointing in Shared Space of Humans and Robots. In *Spatial Cognition*. (cited on Pg. 23, 61, 116)

Restivo, S. (2001). Bringing up and booting up: Social theory and the emergence of socially intelligent robots. In *International Conference on Systems, Man, and Cybernetics* (pp. 2110–2117). IEEE, doi:10.1109/ICSMC.2001.972867. (cited on Pg. 7)

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The Art of Conversation Is Coordination: Common Ground and the Coupling of Eye Movements During Dialogue. *Psychological Science*, 18(5), 407–413, doi:10.1111/j.1467-9280.2007.01914.x. (cited on Pg. 9)

Rickert, M., Foster, M. E., Giuliani, M., By, T., & Panin, G. (2007). Integrating Language, Vision and Action for Human Robot Dialog Systems. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Ambient Interaction*, volume 4555 of *Lecture Notes in Computer*

Science (pp. 987–995). Springer Berlin / Heidelberg, doi:10.1007/978-3-540-73281-5_108, isbn:978-3-540-73280-8. (cited on Pg. 1)

Riek, L. D., Rabinowitch, T.-C., Bremner, P., Pipe, A. G., Fraser, M., & Robinson, P. (2010). Cooperative gestures: Effective signaling for humanoid robots. In *International Conference on Human-Robot Interaction* (pp. 61–68). ACM/IEEE, doi:10.1109/HRI.2010.5453266. (cited on Pg. 14, 37)

Riether, N., Hegel, F., Wrede, B., & Horstmann, G. (2012). Social Facilitation with Social Robots? In *International Conference on Human-Robot Interaction* (pp. 41–47). ACM/IEEE, doi:10.1145/2157689.2157697. (cited on Pg. 3)

Salem, M., Kopp, S., & Joublin, F. (2013a). Closing the Loop: Towards Tightly Synchronized Robot Gesture and Speech. In G. Herrmann, M. Pearson, A. Lenz, P. Bremner, A. Spiers, & U. Leonards (Eds.), *Social Robotics*, volume 8239 of *Lecture Notes in Computer Science* (pp. 381–391). Springer International Publishing, doi:10.1007/978-3-319-02675-6_38, isbn:978-3-319-02674-9. (cited on Pg. 8)

Salem, M., Ziadee, M., & Sakr, M. (2013b). Effects of Politeness and Interaction Context on Perception and Experience of HRI. In G. Herrmann, M. Pearson, A. Lenz, P. Bremner, A. Spiers, & U. Leonards (Eds.), *Social Robotics*, volume 8239 of *Lecture Notes in Computer Science* (pp. 531–541). Springer International Publishing, doi:10.1007/978-3-319-02675-6_53, isbn:978-3-319-02674-9. (cited on Pg. 49, 50)

Schegloff, E. A. (2002). Opening Sequencing. In J. E. Katz & M. A. Aakhus (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 326–385). Cambridge University Press, isbn:0-521-80771-9. (cited on Pg. 21, 31)

Schiller, N. & de Ruiter, J. (2004). Some notes on priming, alignment and self-monitoring. *Behavioral and Brain Sciences*, 27(2), 208–209, doi:10.1017/S0140525X0441005X. (cited on Pg. 9)

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, doi:10.1023/A:1025708916924. (cited on Pg. 57)

Scorolli, C., Miatton, M., Wheaton, L. A., & Borghi, A. M. (2014). I give you a cup, I get a cup: A kinematic study on social intention. *Neuropsychologia*, 57, 196–204, doi:10.1016/j.neuropsychologia.2014.03.006. (cited on Pg. 22)

Sheridan, T. B. (1997). Eight ultimate challenges of human-robot communication. In *International Workshop on Robot and Human Communication* (pp. 9–14). IEEE, doi:10.1109/ROMAN.1997.646944. (cited on Pg. 3, 9)

Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2010). A larger audience, please!: Encouraging people to listen to a guide robot. In *International Conference on Human-Robot Interaction* (pp. 31–38). ACM/IEEE, doi:10.1145/1734454.1734469. (cited on Pg. 22)

Simmons, R., et al. (2011). Believable Robot Characters. *AI Magazine*, 32(4), 39–52, doi:10.1609/aimag.v32i4.2383. (cited on Pg. 13)

Stubbs, K., Wettergreen, D., Hinds, P. J., Clark, H., & Wilkes-gibbs, D. (2007). Autonomy and common ground in human-robot interaction: A field study. *Intelligent Systems*, 22(2), 42–50, doi:10.1109/MIS.2007.21. (cited on Pg. 2)

Sundstrom, E. & Altman, I. (1976). Interpersonal Relationships and Personal Space: Research Review and Theoretical Model. *Human Ecology*, 4(1), 47–67, doi:10.1007/BF01531456. (cited on Pg. 20)

Swadzba, A., Beuter, N., Wachsmuth, S., & Kummert, F. (2010). Using Articulated Scene Models for Dynamic 3D Scene Analysis in Vista Spaces. *3D Research*, 1(3), 20:1–20:13, doi:10.1007/3DRes.03(2010)04. (cited on Pg. 55)

Swadzba, A., Wachsmuth, S., Vorwerg, C., & Rickheit, G. (2009). A Computational Model for the Alignment of Hierarchical Scene Representations in Human-Robot Interaction. In *International Joint Conference on Artificial Ingelligence* (pp. 1857–1863). ACM Press. `http://dl.acm.org/citation.cfm?id=1661743`. (cited on Pg. 11)

Syrdal, D. S., Dautenhahn, K., Walters, M. L., & Koay, K. L. (2008). Sharing Spaces with Robots in a Home Scenario - Anthropomorphic Attributions and their Effect on Proxemic Expectations and Evaluations in a Live HRI Trial. In *Symposium AI in Eldercare: New Solutions to Old Problems* (pp. 116–123). AAAI. `http://www.aaai.org/Papers/Symposia/Fall/2008/FS-08-02/FS08-02-018.pdf`. (cited on Pg. 20)

Syrdal, D. S., Lee Koay, K., Walters, M. L., & Dautenhahn, K. (2007). A personalized robot companion? - The role of individual differences on spatial preferences in HRI scenarios. In *International Symposium on Robot and Human Interactive Communication* (pp. 1143–1148). IEEE, doi:10.1109/ROMAN.2007.4415252. (cited on Pg. 20)

Takayama, L. & Pantofaru, C. (2009). Influences on Proxemic Behaviors in Human-Robot Interaction. In *International Conference on Intelligent Robots and Systems* (pp. 5495–5502). IEEE/RSJ, doi:10.1109/IROS.2009.5354145. (cited on Pg. 20, 37)

Tapus, A. & Mataric, M. J. (2009). The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In *International Conference on Rehabilitation Robotics* (pp. 924–929). IEEE, doi:10.1109/ICORR.2009.5209501. (cited on Pg. 3)

Teneggi, C., Canzoneri, E., di Pellegrino, G., & Serino, A. (2013). Social modulation of peripersonal space boundaries. *Current Biology*, 23(5), 406–411, doi:10.1016/j.cub.2013.01.043. (cited on Pg. 22)

Torta, E., Cuijpers, R. H., Juola, J. F., & van der Pol, D. (2011). Design of Robust Robotic Proxemic Behaviour. In B. Mutlu, C. Bartneck, J. Ham, V. Evers, & T. Kanda (Eds.), *Social Robotics*, volume 7072 of *Lecture Notes in Computer Science* (pp. 21–30). Springer Berlin / Heidelberg, doi:10.1007/978-3-642-25504-5_3, isbn:978-3-642-25503-8. (cited on Pg. 20)

Triplett, N. (1898). The Dynamogenic Factors in Pacemaking and Competition. *The American Journal of Psychology*, 9(4), 507–533, doi:10.2307/1412188. (cited on Pg. 3)

Tsagarakis, N. G., et al. (2007). iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research. *Advanced Robotics*, 21(10), 1151–1175, doi:10.1163/156855307781389419. (cited on Pg. 54)

van Oosterhout, T. & Visser, A. (2008). A visual method for robot proxemics measurements. In *Workshop "Metrics for Human-Robot Interaction" at International Conference on Human-Robot Interaction* (pp. 61–68). University of Hertfordshire. `http://hdl.handle.net/11245/2.58775`. (cited on Pg. 20)

Viola, P. & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *International Conference on Computer Vision and Pattern Recognition* (pp. 511–518). IEEE, doi:10.1109/CVPR.2001.990517. (cited on Pg. 55, 67)

Wachsmuth, I., de Ruiter, J., Jaecks, P., & Kopp, S., Eds. (2013). *Alignment in Communication: Towards a new theory of communication*, volume 6 of *Advances in Interaction Studies*. John Benjamins Publishing, isbn:9789027204608. (cited on Pg. 10)

Wada, K. & Shibata, T. (2007). Living With Seal Robots - Its Sociopsychological and Physiological Influences on the Elderly at a Care House. *Transactions on Robotics*, 23(5), 972–980, doi:10.1109/TRO.2007.906261. (cited on Pg. 1)

Walters, M. L., Koay, K. L., Syrdal, D. S., Dautenhahn, K., & te Boekhorst, R. (2009). Preferences and Perceptions of Robot Appearance and Embodiment in Human-Robot Interaction Trials. In *Symposium "New Frontiers in Human-Robot Interaction" at Artificial Intelligence and Simulation of Behaviour Convention* (pp. 136–143). `http://hdl.handle.net/2299/9642`. (cited on Pg. 54)

Watson, M. O. (1972). Conflicts and Directions in Proxemic Research. *The Journal of Communication*, 22(4), 443–459, doi:10.1111/j.1460-2466.1972.tb00167.x. (cited on Pg. 17)

Williams, M.-A., Abidi, S., Gärdenfors, P., Wang, X., Kuipers, B., & Johnston, B. (2013). Interpreting Robot Pointing Behavior. In G. Herrmann, M. Pearson, A. Lenz, P. Bremner, A. Spiers, & U. Leonards

(Eds.), *Social Robotics*, volume 8239 of *Lecture Notes in Computer Science* (pp. 148–159). Springer International Publishing, doi:10.1007/978-3-319-02675-6_15, isbn:978-3-319-02674-9. (cited on Pg. 61)

Wrede, B., Kopp, S., Rohlfing, K., Lohse, M., & Muhl, C. (2010). Appropriate feedback in asymmetric interactions. *Journal of Pragmatics*, 42(9), 2369–2384, doi:10.1016/j.pragma.2010.01.003. (cited on Pg. 8)

Wrede, B., Rohlfing, K. J., Hanheide, M., & Sagerer, G. (2009). Towards Learning by Interacting. In B. Sendhoff, E. Körner, O. Sporns, H. Ritter, & K. Doya (Eds.), *Creating Brain-Like Intelligence*, volume 5436 of *Lecture Notes in Computer Science* (pp. 139–150). Springer Berlin / Heidelberg, doi:10.1007/978-3-642-00616-6_8, isbn:978-3-642-00615-9. (cited on Pg. 14)

Wrede, S., Fritsch, J., Bauckhage, C., & Sagerer, G. (2004a). An XML Based Framework for Cognitive Vision Architectures. In *International Conference on Pattern Recognition* (pp. 757–760). IEEE, doi:10.1109/ICPR.2004.1334304. (cited on Pg. 55)

Wrede, S., Hanheide, M., Bauckhage, C., & Sagerer, G. (2004b). An Active Memory as a Model for Information Fusion. In *International Conference on Information Fusion* (pp. 198–205). IEEE. `http://www.fusion2004.foi.se/papers/IF04-0198.pdf`. (cited on Pg. 55)

Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2010). A model of proximity control for information-presenting robots. *Transactions on Robotics*, 26(1), 187–195, doi:10.1109/TRO.2009.2035747. (cited on Pg. 21)

Yonezawa, T., Yamazoe, H., Utsumi, A., & Abe, S. (2007). Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking. In *International Conference on Multimodal Interfaces* (pp. 140–145). New York, New York, USA ACM Press, doi:10.1145/1322192.1322218. (cited on Pg. 13)

Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., & Miyamoto, T. (2006). The effects of responsive eye movement and blinking behavior in a communication robot. In *International Conference on Intelligent Robots and Systems* (pp. 4564–4569). IEEE/RSJ, doi:10.1109/IROS.2006.282160. (cited on Pg. 13)

Yu, C., Scheutz, M., & Schermerhorn, P. (2010). Investigating Multimodal Real-Time Patterns of Joint Attention in an HRI Word Learning Task. In *International Conference on Human-Robot Interaction* (pp. 309–316). ACM/IEEE, doi:10.1109/HRI.2010.5453181. (cited on Pg. 14)

Zwaan, R. A. & Radvansky, G. A. (1998). Situation Models in Language Comprehension and Memory. *Psychological Bulletin*, 123(2), 162–185, doi:10.1037/0033-2909.123.2.162. (cited on Pg. 9)

A. Questionnaire Peripersonal Space

0%

Einverständniserklärung

Hiermit erkläre ich mich damit einverstanden, dass im Rahmen dieser Videostudie zum Verhältnis zwischen Mensch und Roboter u.a. **personenbezogene Daten** von mir erfasst werden. Diese Daten werden **vertraulich** und ausschließlich zur **wissenschaftlichen Auswertung** der Studie verwendet. **Anonymisierte Ergebnisse** der Studie können auf wissenschaftlichen Veranstaltungen wie Konferenzen oder Kolloquien gezeigt werden.
Mir ist bewusst, dass meine Antworten mit personenbezogenen Daten (z.B. Alter, Geschlecht, Studiengang oder Beruf) in Verbindung gebracht werden können.

Die Einwilligung über die Verwendung dieser Angaben ist jederzeit widerruflich.
Im Falle des Widerrufs dürfen personenbezogene Daten zukünftig nicht mehr für die oben genannten Zwecke verwendet werden und sind unverzüglich zu löschen.
Soweit die Einwilligung nicht widerrufen wird, gilt sie zeitlich unbeschränkt.
Die Einwilligung ist freiwillig. Aus der Verweigerung der Einwilligung oder ihrem Widerruf entstehen keine Nachteile.

Privacy statement

I hereby agree that the data collected during this study on the relationship between humans and robots includes **personal data**. This data is handled **confidentially** and will only be used for the **scientific analysis** of this study. **Anonymous results** may be shown at scientific events like conferences or colloquiums.
I am aware that my answers can be related to personal data of mine (e.g. age, sex, field of study).

This agreement can be revoked at any time.
In this case, personal data may not be used anymore for the above purposes - they have to be deleted immediately. If the agreement is not revoked, it is valid indefinately.
This agreement is voluntarily. There are no disadvantages caused by the agreement or its revokation.

☐ Ich bin ausreichend über den Datenschutz informiert worden und bin damit einverstanden, dass meine Daten zu wissenschaftlichen Zwecken genutzt werden dürfen.
I have been informed sufficiently about the privacy policy and I agree to the scientific usage of my data.

0%

Willkommen zur Roboter Interaktions-Studie!

Auf den folgenden Seiten haben Sie die Möglichkeit bei einer **Studie** zur **Mensch-Roboter Interaktion** teilzunehmen. Dabei werden Sie gebeten einige Fragen zu beantworten, mit denen Sie das **Verhalten eines Roboters bewerten**. Diese Fragen beziehen sich auf Bilder und Videos, die während der Befragung angezeigt werden.
Die Dauer der Studie beträgt etwa **10 Minuten**, wobei Sie Ihren Fortschritt jederzeit oben auf der Seite verfolgen können.

Sobald Sie alle Fragen auf einer Seite beantwortet haben, klicken Sie bitte auf "OK" am unteren rechten Rand der Seite. So gelangen Sie auf die nächste Seite. Falls Sie im nachhinein noch Antworten auf einer vorherigen Seite verändern möchten, ist dies möglich indem Sie auf "Zurück" klicken. Es wird Ihnen die letzte Seite noch einmal angezeigt.

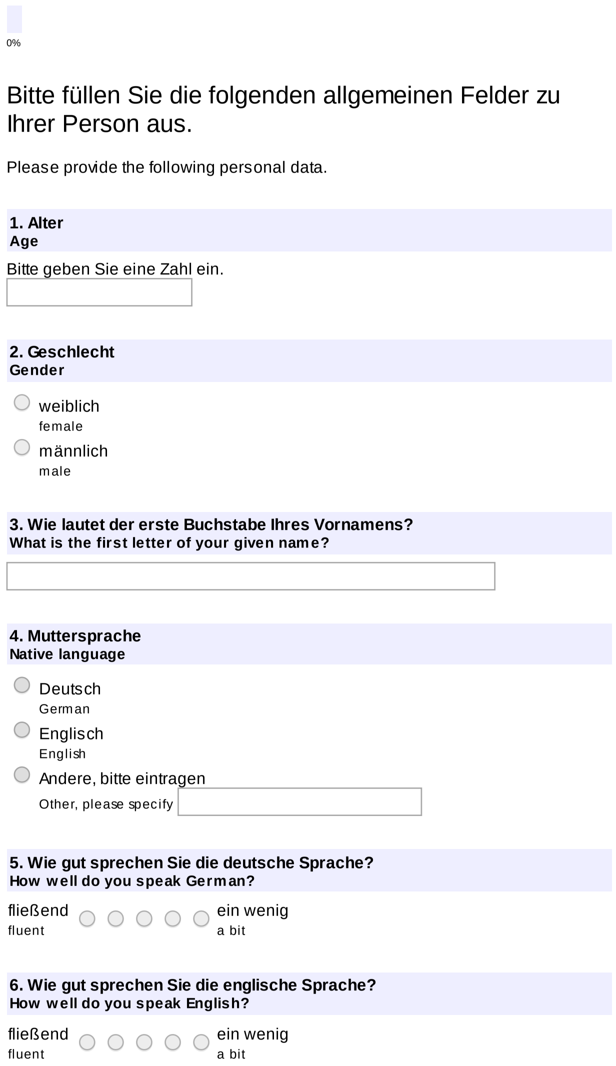
Achtung: Die Umfrage ist nur mit einem aktuellen Webbrowser durchführbar, der Videos anzeigen kann.

Welcome to the robot interaction study!

On the following pages you will have the possibility to contribute to a **study** on **human-robot interaction**. You will be asked to answer a few questions aimed at **judging a robot's behavior**. These questions refer to images and videos that will be played during the questionnaire.
The study will take approximately **10 minutes**. You will be able to see the progress at any time on the top of the page.

As soon as you have answered all the questions on one page, please click the "OK" button at the bottom of that page. After clicking, the next page will be shown. If you want to correct answers from a previous page, you are able to do so by clicking the "Zurück" button. You will be led to the previous page.

Attention: The questionnaire is only usable with a recent web browser that is capable of displaying videos.



0%

Bitte füllen Sie die folgenden allgemeinen Felder zu Ihrer Person aus.

Please provide the following personal data.

1. Alter
Age

Bitte geben Sie eine Zahl ein.

2. Geschlecht
Gender

○ weiblich
female

○ männlich
male

3. Wie lautet der erste Buchstabe Ihres Vornamens?
What is the first letter of your given name?

4. Muttersprache
Native language

○ Deutsch
German

○ Englisch
English

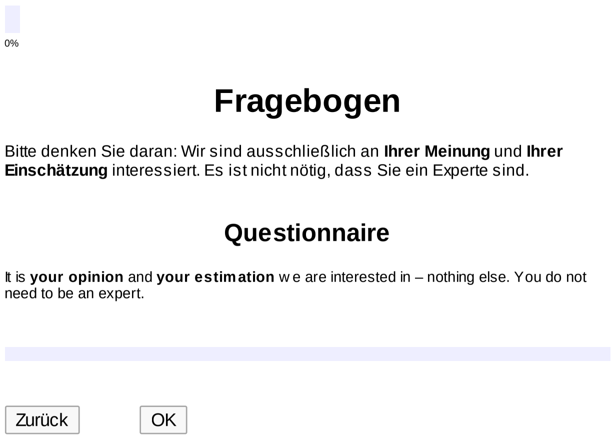
○ Andere, bitte eintragen
Other, please specify

5. Wie gut sprechen Sie die deutsche Sprache?
How well do you speak German?

fließend ○ ○ ○ ○ ○ ein wenig
fluent a bit

6. Wie gut sprechen Sie die englische Sprache?
How well do you speak English?

fließend ○ ○ ○ ○ ○ ein wenig
fluent a bit



0%

Fragebogen

Bitte denken Sie daran: Wir sind ausschließlich an **Ihrer Meinung** und **Ihrer Einschätzung** interessiert. Es ist nicht nötig, dass Sie ein Experte sind.

Questionnaire

It is **your opinion** and **your estimation** we are interested in – nothing else. You do not need to be an expert.

28%

Beruf und Arbeitsfeld.

Occupation and field of interest.

8. Wie lautet Ihre Berufsbezeichnung?
What is your occupation?

○ Student
Student
○ Angestellter
Employee
○ Selbstständig
Self employed
○ Andere, bitte eintragen
Other, please specify ☐

Zurück OK

---

7. Mit welcher Hand sind Sie am geschicktesten?
What hand do you feel most dexterous with?

○ Mit der linken Hand
With the left hand
○ Mit der rechten Hand
With the right hand
○ Mit beiden Händen exakt gleich
Both hands equally

Zurück OK

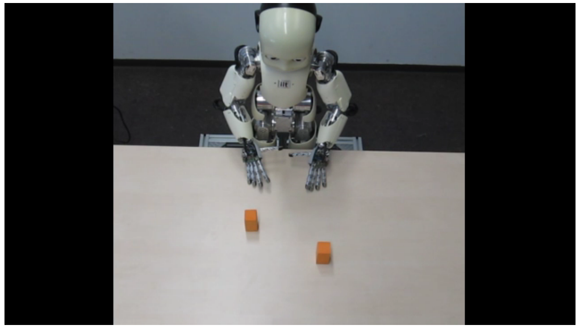
56%

Szene I

Weiter unten auf der Seite sehen Sie die Momentaufnahme einer Interaktion zwischen Mensch und Roboter aus der Perspektive des Menschen. Beide haben die Möglichkeit, ein Objekt in die Hand zu nehmen und anzusehen. In dieser Szene ist zunächst der Roboter an der Reihe, ein Objekt auszuwählen.

Scene I

Further down on this page, you can see a snapshot of an interaction between a human and a robot from the human's perspective. Both have the possibility to pick an object and inspect it. In this scene, it is the robot's turn to choose an object.



15. Welches Objekt soll der Roboter Ihrer Meinung nach als nächstes greifen?
Which object should the robot grasp next in your opinion?

○ Das Objekt **links** ←
The object on the **left**

○ Das Objekt **rechts** →
The object on the **right**

16. Warum sollte der Roboter dieses Objekt wählen?
Why should the robot choose this object?

Zurück OK

44%

Erfahrung mit Computern und Robotern.

Experience with robots and computers.

12. Bitte bewerten Sie ihre Erfahrung mit *Computern* im Allgemeinen.
Please rate your general level of experience with *computers*.

sehr erfahren ○ ○ ○ ○ ○ **nicht** erfahren
very experienced **not** experienced

13. Bitte nennen Sie uns *alle* Ihnen bekannte Roboter.
Please indicate *all* the robots you know from the list below.

- ☐ Asimo
- ☐ BIRON
- ☐ Flobi
- ☐ Geminoid
- ☐ iCub
- ☐ Kismet
- ☐ Roboter aus Filmen
  Robots from movies
- ☐ Spielzeugroboter
  Toy robots
- ☐ Industrieroboter
  Industrial robots
- ☐ Andere Roboter, bitte angeben
  Other robots, please specify

14. Bitte bewerten Sie ihre Erfahrung mit Robotern im Allgemeinen.
Please rate your level of experience with robots.

sehr erfahren ○ ○ ○ ○ ○ **nicht** erfahren
very experienced **not** experienced

Zurück OK

17. Welches Objekt soll der Roboter Ihrer Meinung nach als nächstes greifen?
Which object should the robot grasp next in your opinion?

○ Das Objekt **links** ←
The object on the **left**

○ Das Objekt **rechts** →
The object on the **right**

18. Warum sollte der Roboter dieses Objekt wählen?
Why should the robot choose this object?

Zurück OK

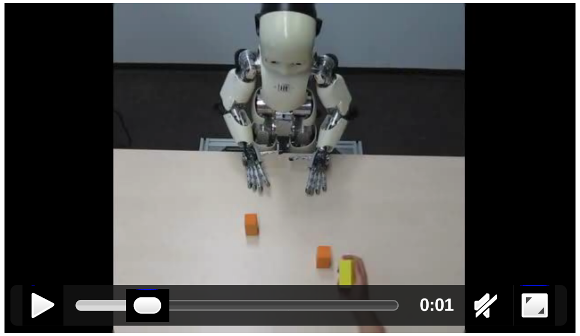
64%

Szene II

Auf dieser Seite sehen Sie einen Video-Ausschnitt aus einer Interaktion zwischen Mensch und Roboter aus der Perspektive des Menschen. Beide haben wieder die Möglichkeit, ein Objekt in die Hand zu nehmen und anzusehen. In dieser Szene Szene sieht sich zuerst der Mensch ein Objekt an, woraufhin wieder der Roboter an der Reihe ist.
Sie haben die Möglichkeit das Video jederzeit und so häufig Sie möchten zu starten oder zu pausieren.

Scene II

On this page, you can see a video scene of an interaction between a human and a robot from the human's perspective. Once more, both have the possibility to pick an object and inspect it. In this scene, the human inspects an object first, then it is again the robot's turn to choose an object.
You can play, pause or restart the video below at any time and as often as you want.



Bitte sehen Sie sich das gesamte Video mindestens einmal an, bevor Sie mit der Beantwortung der Fragen beginnen.
Please watch the whole video at least once before answering any of the questions.

19. Welches Objekt soll der Roboter Ihrer Meinung nach als nächstes greifen?
Which object should the robot grasp next in your opinion?

○ Das Objekt **links** ←
The object on the **left**

○ Das Objekt **rechts** →
The object on the **right**

20. Warum sollte der Roboter dieses Objekt wählen?
Why should the robot choose this object?

Zurück

OK

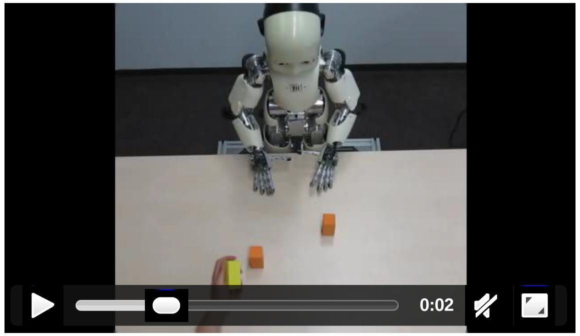
72%

Szene III

Auf dieser Seite sehen Sie ebenfalls einen Video-Ausschnitt aus einer Interaktion zwischen Mensch und Roboter aus der Perspektive des Menschen. Beide haben wieder die Möglichkeit, ein Objekt in die Hand zu nehmen und anzusehen. Wie in der letzten Szene sieht sich zuerst der Mensch ein Objekt an, woraufhin wieder der Roboter an der Reihe ist.
Sie haben die Möglichkeit das Video jederzeit und so häufig Sie möchten zu starten oder zu pausieren.

Scene III

On this page, you can again see a video scene of an interaction between a human and a robot from the human's perspective. Once more, both have the possibility to pick an object and inspect it. As in the last scene, the human inspects an object first, then it is again the robot's turn to choose an object.
You can play, pause or restart the video below at any time and as often as you want.



Bitte sehen Sie sich das gesamte Video mindestens einmal an, bevor Sie mit der Beantwortung der Fragen beginnen.
Please watch the whole video at least once before answering any of the questions.

96%

Ende

Falls Sie noch Anmerkungen, Kommentare oder Kritik zu dieser Studie haben, dürfen Sie diese jetzt in dem unten stehenden Feld anbringen.

The End

If you have any comments or criticism about this study, then please tell us using the field below.

25. Kommentare:
Comments:

☐

Zurück Abschicken

80%

Über den Roboter

Auf dieser Seite finden Sie einige Fragen zur Bewertung des auf den vorherigen Seiten gezeigten Roboters.

About the Robot

On this page, you can find some questions regarding the evaluation of the robot from the last pages.

21. Glauben Sie, dass der Roboter die Geste des Menschen wahrnimmt?
Do you think the robot recognizes the human's gesture?

○ Ja
Yes
○ Nein
No

22. Glauben Sie, dass der Roboter technisch in der Lage ist, eines der Objekte zu greifen?
Do you think the robot is physically capable of grasping one of the objects?

○ Ja
Yes
○ Nein
No

23. Glauben Sie, dass die Geste des Menschen die Auswahl des Roboters beeinflusst?
Do you think the human's gesture influences the robot's choice?

○ Ja
Yes
○ Nein
No

24. Glauben Sie, dass der Abstand der Objekte zum Roboter die Auswahl des Roboters beeinflusst?
Do you think the distance between object and robot influences the robot's choice?

○ Ja
Yes
○ Nein
No

B. Receptionist Grammar

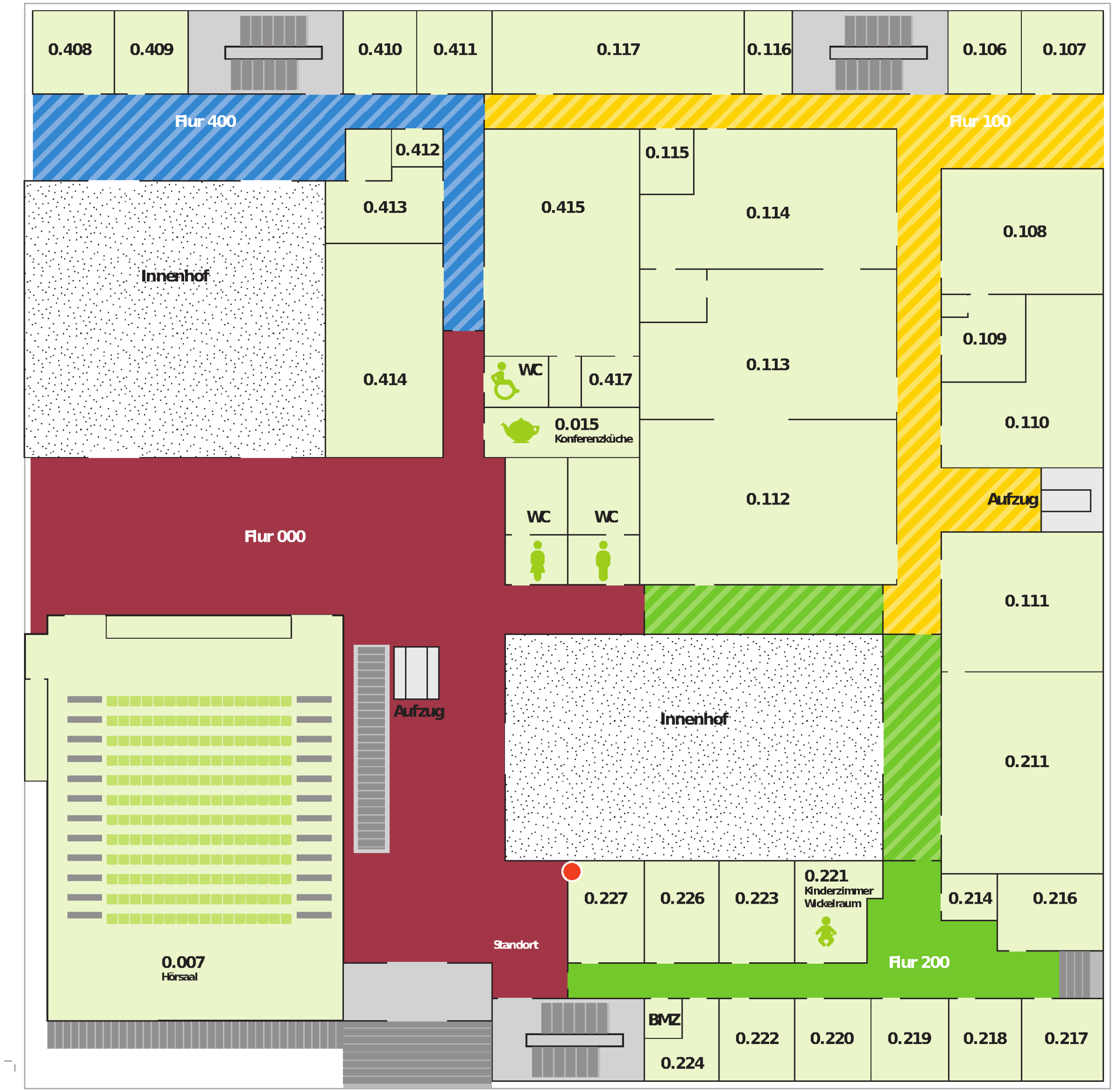
```
$$S:            $greet
                        | $reset
                        | $end
                        | $help
                        | $map ;

$robot:         icub | robby | robot ;

$greet:         hello
                        | hello again
                        | $robot hello
                        | hello $robot ;

$reset:         $restart
                        | $restart $robot
                        | $robot $restart ;

$restart:       reset
                        | start over ;

$end:           good bye
                        | bye bye
                        | bye ;

$help:          $robot help me
                        | help me $robot
                        | help me
                        | what can you do
                        | $robot what can you do
                        | what can you do $robot ;

$roomname:      disabled's restroom
                        | women's restroom
                        | men's restroom
                        | restroom
                        | first courtyard
                        | second courtyard
                        | stairs
                        | elevator
                        | secondary elevator
                        | b m z
                        | entrance
                        | exit ;

$roomdescription: kitchen
                        | secretary
                        | central lab
                        | bimanual interaction lab
                        | electro studio
                        | mechanics studio
                        | ambient intelligence lab
                        | augmented reality lab
                        | interactive media lab
                        | cis lab
                        | industry transfer lab
                        | humanoids lab
                        | cognitronics teleworkbench
                        | biomechanics lab
                        | auditorium
                        | fire alarm control
                        | nursery
                        | smaller courtyard
                        | bigger courtyard ;

$roomperson:    sven wachsmuth
                        | anita adamczyk;


$number:        oh | zero | one | two | three | four | five | six |
seven | eight | nine | ten ;

$please:        please ;

$what:          what's | what is ;

$where:         where's | where is ;

$this:          this ;

$room:          room
                        | room number
                        | the room
                        | the room number ;

$roomnumber:    $number $number $number ;

$wherepl:       where are ;

$plural:        you
                        | we
                        | the restrooms
                        | the stairs ;

$map:           $what $this
                        | $wherepl $plural
                        | $where $this
                        | $where $robot
                        | $where the $robot
                        | $where $roomnumber
                        | $where $room $roomnumber
                        | $where the $roomname
                        | $where the $roomdescription
                        | $where the office of $roomperson;
```

C. Receptionist Utterances

```
Blocked by human in simple strategy (adv.):
Sorry, you're in my way.

Joints Disabled/disconnected:
Sorry, I am not allowed to use my arms.

Timeout reaching safety position:
Sorry, could not act safely, aborting.


===== Prompting (adv.)
Attempt #1:
This/These...

Attempt #2:
This is/These are...

Attempt #3:
Sorry...

Blocked by human and maximum prompts/time reached:
Sorry, I could not point towards the map.


===== Pointing Towards Real Location
Standard:
You can find it/them in this direction (behind me).

Pointing towards map unsuccessful:
You can find <room>, <annotation> in this direction (behind me).

Disabled:
Sorry, I cannot show <room> to you.


===== Post Explaination
Can I tell you something else?
May I show you another location?
Do you want to know another location?
```

```
===== Greeting
Hello. What can I do for you?
Hello, nice to see you.
Hello, I am I Cub. How can I help?


===== Farewell
See you soon.
See you.


===== Help
Human-triggered:
I can explain you how to find your way with the help of <map>.

Robot-triggered (adv.):
I am I cub, a receptionist robot. I can explain you how to find your
way with the help of <map>. What would you like to know?


===== Error handling
Internal Error (joints disabled):
Sorry, I cannot show <room> to you.

Item not found:
Sorry, I cannot find <room>.
Sorry, I cannot find <description>.
Sorry, I cannot see what you are pointing at.

Internal Error:
Sorry, I could not point towards the map.


===== Avoidance (adv.)
Please don't come too close.


===== Reset
Okay, let's start over.
Alright, I'm resetting myself.
Okay, the system will be reset.
System reset.
Reloading configuration.


===== Pointing Towards Map
Standard:
This is/these are <room>, <annotation>.

Without annotation:
This is/these are <room>.

Current location:
This is <room>. We are currently here.

Robot:
This is where I am.
```

D. Floor Plan



Erdgeschoss

Flure — **geschlossen**

- 🟥 Flur 000
- 🟨 Flur 100
- 🟩 Flur 200
- 🟧 Flur 300
- 🟦 Flur 400

▨ ▨ ▨

Geschäftsstelle	
Anita Adamczyk	0.226
Sekretariat	0.227

CITEC - Zentrallabor- Bereich	
Dr.- Ing. Sven Wachsmuth	0.408

Labore	
Zentrallabor | CITEC	0.113
Zentrallabor | CITEC	0.114
Bimanual Interaction Lab | CITEC	0.112
Elektro- Werkstatt | CITEC	0.110
Mechanik- Werkstatt | CITEC	0.413
Ambient Intelligence Lab	0.108
Augmented Reality Lab	0.109
Interactive Media Lab	0.117
CIS Lab	0.116
Industrie- Transfer- Lab | CoR- Lab	0.211
Humanoids Lab | CoR- Lab	0.111
Kognitronik Telewerkbank	0.414
Biomechanics Lab	0.415

```
 <!--  humanoids/map.xml  -->
▼<map>
   <defaults ref="map" type="xyz" unit="mm"/>
 ▶<mapdetails>...</mapdetails>
 ▼<mapcontent>
   ▼<objects>
     ▼<object name="iCub" type="robot">
        <annotation>me on the map</annotation>
        <coordinates ref="real" unit="m" d1="-.1" d2="0" d3="-.1"
        r1=".2" r2="1" r3=".2"/>
      </object>
     ▼<object name="room-entry" type="fixed">
        <annotation>the door</annotation>
        <coordinates ref="real" unit="m" d1="3" d2="1.2" d3="5.7"
        r1=".2" r2=".2" r3=".2"/>
      </object>
     ▼<object name="0.107" type="building">
        <coordinates ref="map" unit="mm" d1="0" d3="0" d2="0" r1="12"
        r3="13" r2="3"/>
      </object>
     ▼<object name="0.106" type="building">
        <coordinates ref="map" unit="mm" d1="12" d3="0" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="the stairs north east" type="fixed">
        <coordinates ref="map" unit="mm" d1="24" d3="0" d2="-3"
        r1="24" r3="13" r2="9"/>
      </object>
     ▼<object name="0.116" type="building">
        <annotation>the cis lab</annotation>
        <coordinates ref="map" unit="mm" d1="48" d3="0" d2="0" r1="7"
        r3="13" r2="3"/>
      </object>
     ▼<object name="0.117" type="building">
        <annotation>the interactive media lab</annotation>
        <coordinates ref="map" unit="mm" d1="55" d3="0" d2="0"
        r1="39" r3="13" r2="3"/>
      </object>
     ▼<object name="0.411" type="building">
        <coordinates ref="map" unit="mm" d1="94" d3="0" d2="0"
        r1="11" r3="13" r2="3"/>
      </object>
     ▼<object name="0.410" type="building">
        <coordinates ref="map" unit="mm" d1="105" d3="0" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="the stairs north west" type="fixed">
        <coordinates ref="map" unit="mm" d1="117" d3="0" d2="-3"
        r1="23" r3="13" r2="9"/>
      </object>
     ▼<object name="0.409" type="building">
        <coordinates ref="map" unit="mm" d1="140" d3="0" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.408" type="building">
        <annotation>the office of Sven Wachsmuth</annotation>
        <coordinates ref="map" unit="mm" d1="152" d3="0" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.108" type="building">
        <annotation>the ambient intelligence lab</annotation>
        <coordinates ref="map" unit="mm" d1="0" d3="24" d2="0"
        r1="25" r3="20" r2="3"/>
      </object>
     ▼<object name="0.109" type="building">
        <annotation>the ambient reality lab</annotation>
        <coordinates ref="map" unit="mm" d1="12" d3="44" d2="0"
        r1="13" r3="13" r2="3"/>
      </object>
     ▼<object name="0.110" type="building">
        <annotation>the electro studio</annotation>
        <coordinates ref="map" unit="mm" d1="0" d3="44" d2="0"
        r1="25" r3="26" r2="3"/>
      </object>
     ▼<object name="the secondary elevator" type="fixed">
        <coordinates ref="map" unit="mm" d1="0" d3="70" d2="-3"
        r1="9" r3="10" r2="9"/>
      </object>
     ▼<object name="0.111" type="building">
        <annotation>the humanoids lab</annotation>
        <coordinates ref="map" unit="mm" d1="0" d3="80" d2="0"
        r1="25" r3="21" r2="3"/>
      </object>
     ▼<object name="0.211" type="building">
        <annotation>the industry transfer lab</annotation>
        <coordinates ref="map" unit="mm" d1="0" d3="101" d2="0"
        r1="25" r3="32" r2="3"/>
      </object>
     ▼<object name="0.216" type="building">
        <coordinates ref="map" unit="mm" d1="0" d3="133" d2="0"
        r1="16" r3="11" r2="3"/>
      </object>
     ▼<object name="0.214" type="building">
        <coordinates ref="map" unit="mm" d1="16" d3="133" d2="0"
        r1="9" r3="6" r2="3"/>
      </object>
     ▼<object name="0.217" type="building">
        <coordinates ref="map" unit="mm" d1="0" d3="152" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.218" type="building">
        <coordinates ref="map" unit="mm" d1="12" d3="152" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.219" type="building">
        <coordinates ref="map" unit="mm" d1="24" d3="152" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.220" type="building">
        <coordinates ref="map" unit="mm" d1="36" d3="152" d2="0"
        r1="12" r3="13" r2="3"/>
      </object>
     ▼<object name="0.222" type="building">
        <coordinates ref="map" unit="mm" d1="48" d3="152" d2="0"
        r1="11" r3="13" r2="3"/>
```

```
    <coordinates ref="map" unit="mm" d1="85" d3="53" d2="0"
    r1="10" r3="8" r2="3"/>
  </object>
▼<object name="0.015" type="building">
    <annotation>the kitchen</annotation>
    <coordinates ref="map" unit="mm" d1="71" d3="61" d2="0"
    r1="24" r3="7" r2="3"/>
  </object>
▼<object name="0.417" type="building">
    <coordinates ref="map" unit="mm" d1="71" d3="53" d2="0"
    r1="9" r3="8" r2="3"/>
  </object>
▼<object name="the men's restroom" type="building">
    <coordinates ref="map" unit="mm" d1="71" d3="68" d2="0"
    r1="11" r3="20" r2="3"/>
  </object>
▼<object name="the women's restroom" type="building">
    <coordinates ref="map" unit="mm" d1="82" d3="68" d2="0"
    r1="11" r3="20" r2="3"/>
  </object>
▼<object name="the restroom" type="building">
    <coordinates ref="map" unit="mm" d1="71" d3="68" d2="0"
    r1="22" r3="20" r2="3"/>
  </object>
▼<object name="0.412" type="building">
    <coordinates ref="map" unit="mm" d1="101" d3="18" d2="0"
    r1="8" r3="6" r2="3"/>
  </object>
▼<object name="0.413" type="building">
    <annotation>the mechanics studio</annotation>
    <coordinates ref="map" unit="mm" d1="101" d3="24" d2="0"
    r1="18" r3="12" r2="3"/>
  </object>
▼<object name="0.414" type="building">
    <annotation>the cognitronics teleworkbench</annotation>
    <coordinates ref="map" unit="mm" d1="101" d3="36" d2="0"
    r1="18" r3="32" r2="3"/>
  </object>
▼<object name="the second courtyard" type="fixed">
    <annotation>the smaller courtyard</annotation>
    <coordinates ref="map" unit="mm" d1="119" d3="26" d2="0"
    r1="45" r3="42" r2="3"/>
  </object>
▼<object name="the first courtyard" type="fixed">
    <annotation>the bigger courtyard</annotation>
    <coordinates ref="map" unit="mm" d1="34" d3="96" d2="0"
    r1="58" r3="35" r2="3"/>
  </object>
▼<object name="0.007" type="building">
    <annotation>the auditorium</annotation>
    <coordinates ref="map" unit="mm" d1="116" d3="90" d2="0"
    r1="48" r3="69" r2="3"/>
  </object>
▼<object name="the elevator" type="fixed">
    <coordinates ref="map" unit="mm" d1="103" d3="96" d2="-3"
    r1="6" r3="7" r2="9"/>
  </object>
▼<object name="the stairs" type="fixed">
  </object>
▼<object name="0.224" type="building">
    <coordinates ref="map" unit="mm" d1="59" d3="152" d2="0"
    r1="11" r3="13" r2="3"/>
  </object>
▼<object name="the B M Z" type="building">
    <annotation>the fire alarm control</annotation>
    <coordinates ref="map" unit="mm" d1="64" d3="152" d2="0"
    r1="6" r3="6" r2="3"/>
  </object>
▼<object name="the southern stairs" type="fixed">
    <coordinates ref="map" unit="mm" d1="70" d3="152" d2="-3"
    r1="24" r3="13" r2="9"/>
  </object>
▼<object name="0.221" type="building">
    <annotation>the nursery</annotation>
    <coordinates ref="map" unit="mm" d1="33" d3="131" d2="0"
    r1="14" r3="15" r2="3"/>
  </object>
▼<object name="0.223" type="building">
    <coordinates ref="map" unit="mm" d1="47" d3="131" d2="0"
    r1="12" r3="15" r2="3"/>
  </object>
▼<object name="0.226" type="building">
    <annotation>the office of Anita Adamczyk</annotation>
    <coordinates ref="map" unit="mm" d1="59" d3="131" d2="0"
    r1="11" r3="15" r2="3"/>
  </object>
▼<object name="0.227" type="building">
    <annotation>the secretary</annotation>
    <coordinates ref="map" unit="mm" d1="70" d3="131" d2="0"
    r1="12" r3="15" r2="3"/>
  </object>
▼<object name="0.115" type="building">
    <coordinates ref="map" unit="mm" d1="63" d3="18" d2="0"
    r1="8" r3="5" r2="3"/>
  </object>
▼<object name="0.114" type="building">
    <annotation>the central lab</annotation>
    <coordinates ref="map" unit="mm" d1="32" d3="18" d2="0"
    r1="39" r3="22" r2="3"/>
  </object>
▼<object name="0.113" type="building">
    <annotation>the central lab</annotation>
    <coordinates ref="map" unit="mm" d1="32" d3="40" d2="0"
    r1="39" r3="23" r2="3"/>
  </object>
▼<object name="0.112" type="building">
    <annotation>the bimanual interaction lab</annotation>
    <coordinates ref="map" unit="mm" d1="32" d3="63" d2="0"
    r1="39" r3="25" r2="3"/>
  </object>
▼<object name="0.415" type="building">
    <annotation>the biomechanics lab</annotation>
    <coordinates ref="map" unit="mm" d1="71" d3="18" d2="0"
    r1="24" r3="35" r2="3"/>
  </object>
▼<object name="the disabled's restroom" type="building">
```

```
    <coordinates ref="map" unit="mm" d1="110" d3="96" d2="0"
    r1="5" r3="36" r2="6"/>
  </object>
▼<object name="the entrance" type="fixed">
    <coordinates ref="map" unit="mm" d1="95" d3="149" d2="0"
    r1="10" r3="20" r2="3"/>
  </object>
▼<object name="the exit" type="fixed">
    <coordinates ref="map" unit="mm" d1="105" d3="149" d2="0"
    r1="10" r3="20" r2="3"/>
  </object>
 </objects>
</mapcontent>
</map>
```

E. Questionnaire Opening Sequence

Einverständniserklärung

Hiermit erkläre ich mich damit einverstanden, dass im Rahmen dieser Videostudie zum Verhältnis zwischen Mensch und Roboter u.a. **personenbezogene Daten** von mir erfasst werden. Diese Daten werden **vertraulich** und ausschließlich zur **wissenschaftlichen Auswertung** der Studie verwendet. **Anonymisierte Ergebnisse** der Studie können auf wissenschaftlichen Veranstaltungen wie Konferenzen oder Kolloquien gezeigt werden.
Mir ist bewusst, dass meine Antworten mit personenbezogenen Daten (z.B. Alter, Geschlecht, Studiengang oder Beruf) in Verbindung gebracht werden können.

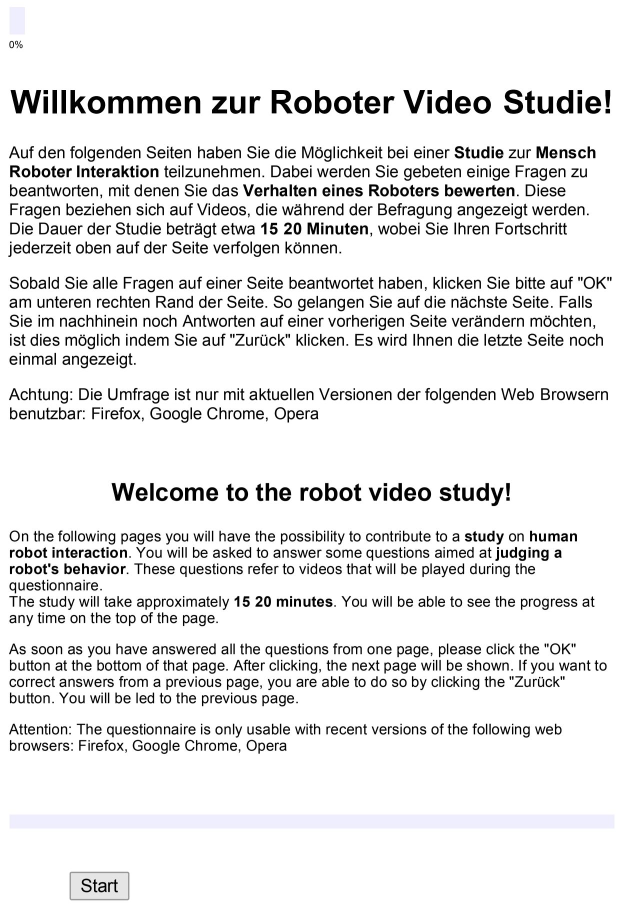
Die Einwilligung über die Verwendung dieser Angaben ist jederzeit widerruflich.
Im Falle des Widerrufs dürfen personenbezogene Daten zukünftig nicht mehr für die oben genannten Zwecke verwendet werden und sind unverzüglich zu löschen.
Soweit die Einwilligung nicht widerrufen wird, gilt sie zeitlich unbeschränkt.
Die Einwilligung ist freiwillig. Aus der Verweigerung der Einwilligung oder ihrem Widerruf entstehen keine Nachteile.

Privacy statement

I hereby agree that among others **personal data** is collected during this study on the relationship between humans and robots. This data is handled **confidentially** and will only be used for the **scientific analysis** of this study. **Anonymous results** may be shown on scientific events like conferences or colloqiums.
I am aware that my answers can be related to personal data of mine (e.g. age, sex, field of study).

This agreement can be revoked at any time.
In this case, personal data may not be used anymore for the above purposes they have to be deleted immediately. If the agreement is not revoked, it is valid infinitely.
This agreement is voluntarily. There are no disadvantages caused by the agreement or its revokation.

☐ Ich bin ausreichend über den Datenschutz informiert worden und bin damit einverstanden, dass meine Daten zu wissenschaftlichen Zwecken genutzt werden dürfen.
I have been informed well about the privacy policy and I agree to the scientific usage of my data.



Willkommen zur Roboter Video Studie!

Auf den folgenden Seiten haben Sie die Möglichkeit bei einer **Studie** zur **Mensch Roboter Interaktion** teilzunehmen. Dabei werden Sie gebeten einige Fragen zu beantworten, mit denen Sie das **Verhalten eines Roboters bewerten**. Diese Fragen beziehen sich auf Videos, die während der Befragung angezeigt werden.
Die Dauer der Studie beträgt etwa **15 20 Minuten**, wobei Sie Ihren Fortschritt jederzeit oben auf der Seite verfolgen können.

Sobald Sie alle Fragen auf einer Seite beantwortet haben, klicken Sie bitte auf "OK" am unteren rechten Rand der Seite. So gelangen Sie auf die nächste Seite. Falls Sie im nachhinein noch Antworten auf einer vorherigen Seite verändern möchten, ist dies möglich indem Sie auf "Zurück" klicken. Es wird Ihnen die letzte Seite noch einmal angezeigt.

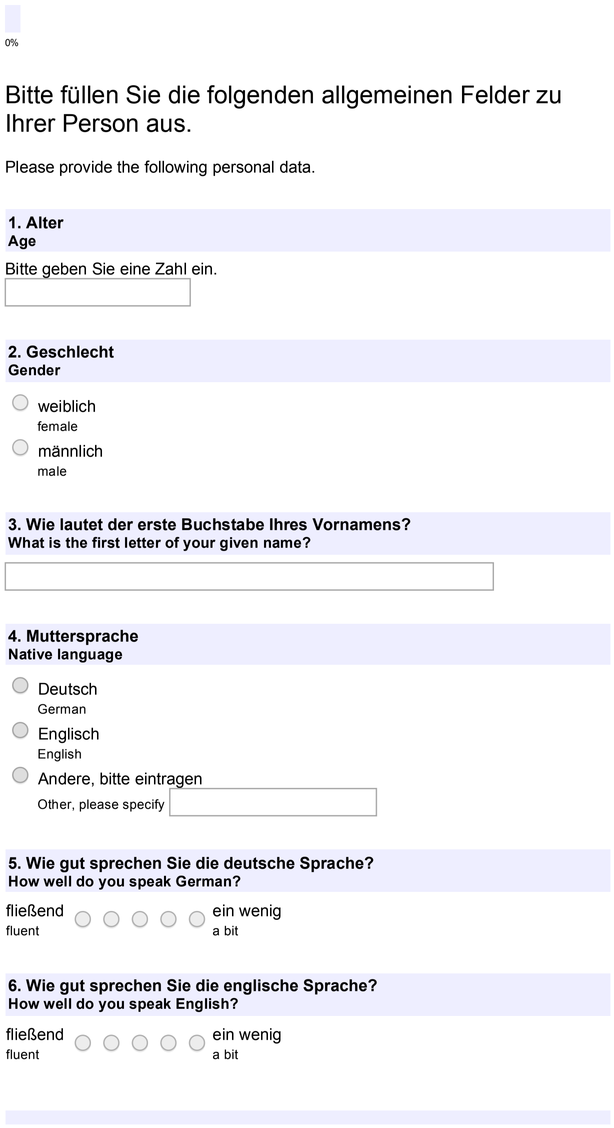
Achtung: Die Umfrage ist nur mit aktuellen Versionen der folgenden Web Browsern benutzbar: Firefox, Google Chrome, Opera

Welcome to the robot video study!

On the following pages you will have the possibility to contribute to a **study** on **human robot interaction**. You will be asked to answer some questions aimed at **judging a robot's behavior**. These questions refer to videos that will be played during the questionnaire.
The study will take approximately **15 20 minutes**. You will be able to see the progress at any time on the top of the page.

As soon as you have answered all the questions from one page, please click the "OK" button at the bottom of that page. After clicking, the next page will be shown. If you want to correct answers from a previous page, you are able to do so by clicking the "Zurück" button. You will be led to the previous page.

Attention: The questionnaire is only usable with recent versions of the following web browsers: Firefox, Google Chrome, Opera



Bitte füllen Sie die folgenden allgemeinen Felder zu Ihrer Person aus.

Please provide the following personal data.

1. Alter
Age

Bitte geben Sie eine Zahl ein.

2. Geschlecht
Gender

○ weiblich
female
○ männlich
male

3. Wie lautet der erste Buchstabe Ihres Vornamens?
What is the first letter of your given name?

4. Muttersprache
Native language

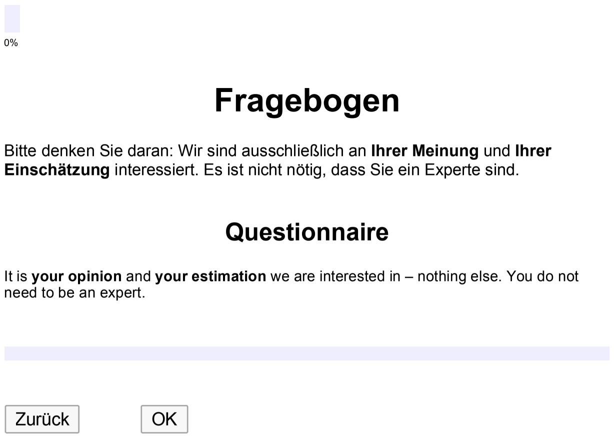
○ Deutsch
German
○ Englisch
English
○ Andere, bitte eintragen
Other, please specify

5. Wie gut sprechen Sie die deutsche Sprache?
How well do you speak German?

fließend fluent ○ ○ ○ ○ ○ ein wenig a bit

6. Wie gut sprechen Sie die englische Sprache?
How well do you speak English?

fließend fluent ○ ○ ○ ○ ○ ein wenig a bit

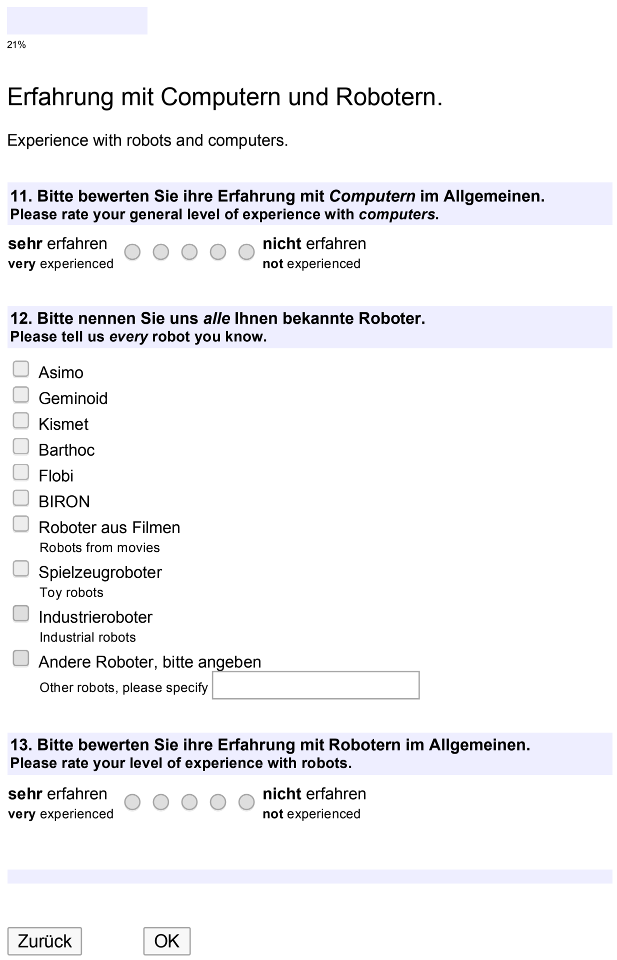


Fragebogen

Bitte denken Sie daran: Wir sind ausschließlich an **Ihrer Meinung** und **Ihrer Einschätzung** interessiert. Es ist nicht nötig, dass Sie ein Experte sind.

Questionnaire

It is **your opinion** and **your estimation** we are interested in – nothing else. You do not need to be an expert.



Erfahrung mit Computern und Robotern.

Experience with robots and computers.

11. Bitte bewerten Sie ihre Erfahrung mit *Computern* im Allgemeinen.
Please rate your general level of experience with *computers*.

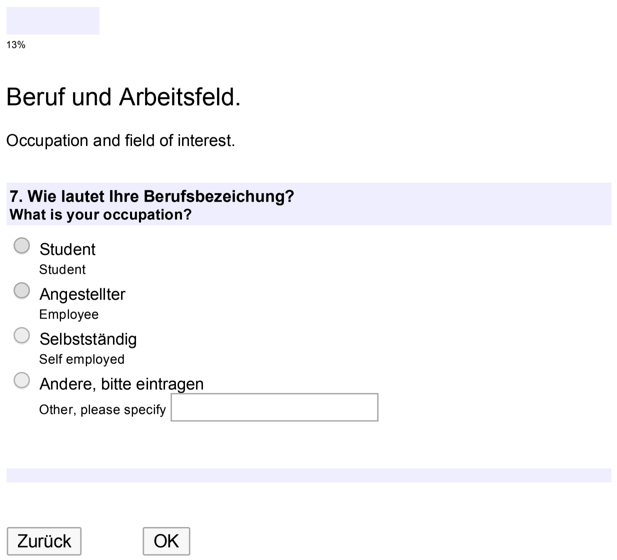
sehr erfahren / **very** experienced ○ ○ ○ ○ ○ **nicht** erfahren / **not** experienced

12. Bitte nennen Sie uns *alle* Ihnen bekannte Roboter.
Please tell us *every* robot you know.

- ☐ Asimo
- ☐ Geminoid
- ☐ Kismet
- ☐ Barthoc
- ☐ Flobi
- ☐ BIRON
- ☐ Roboter aus Filmen / Robots from movies
- ☐ Spielzeugroboter / Toy robots
- ☐ Industrieroboter / Industrial robots
- ☐ Andere Roboter, bitte angeben / Other robots, please specify

13. Bitte bewerten Sie ihre Erfahrung mit Robotern im Allgemeinen.
Please rate your level of experience with robots.

sehr erfahren / **very** experienced ○ ○ ○ ○ ○ **nicht** erfahren / **not** experienced

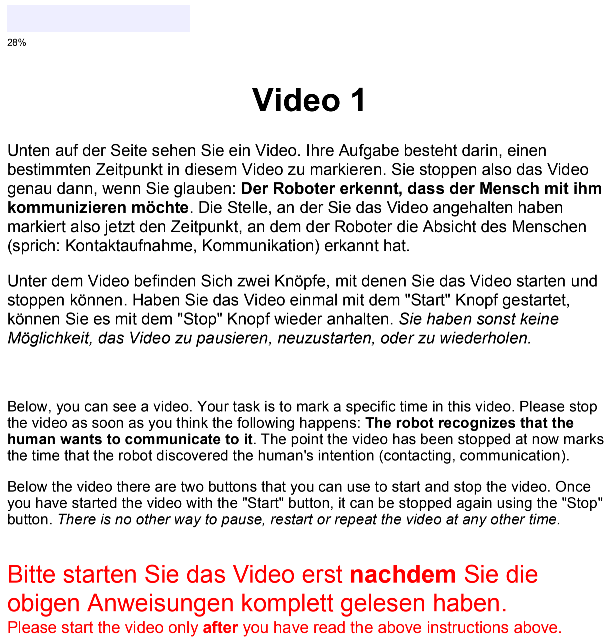


Beruf und Arbeitsfeld.

Occupation and field of interest.

7. Wie lautet Ihre Berufsbezeichung?
What is your occupation?

- ○ Student / Student
- ○ Angestellter / Employee
- ○ Selbstständig / Self employed
- ○ Andere, bitte eintragen / Other, please specify



Video 1

Unten auf der Seite sehen Sie ein Video. Ihre Aufgabe besteht darin, einen bestimmten Zeitpunkt in diesem Video zu markieren. Sie stoppen also das Video genau dann, wenn Sie glauben: **Der Roboter erkennt, dass der Mensch mit ihm kommunizieren möchte**. Die Stelle, an der Sie das Video angehalten haben markiert also jetzt den Zeitpunkt, an dem der Roboter die Absicht des Menschen (sprich: Kontaktaufnahme, Kommunikation) erkannt hat.

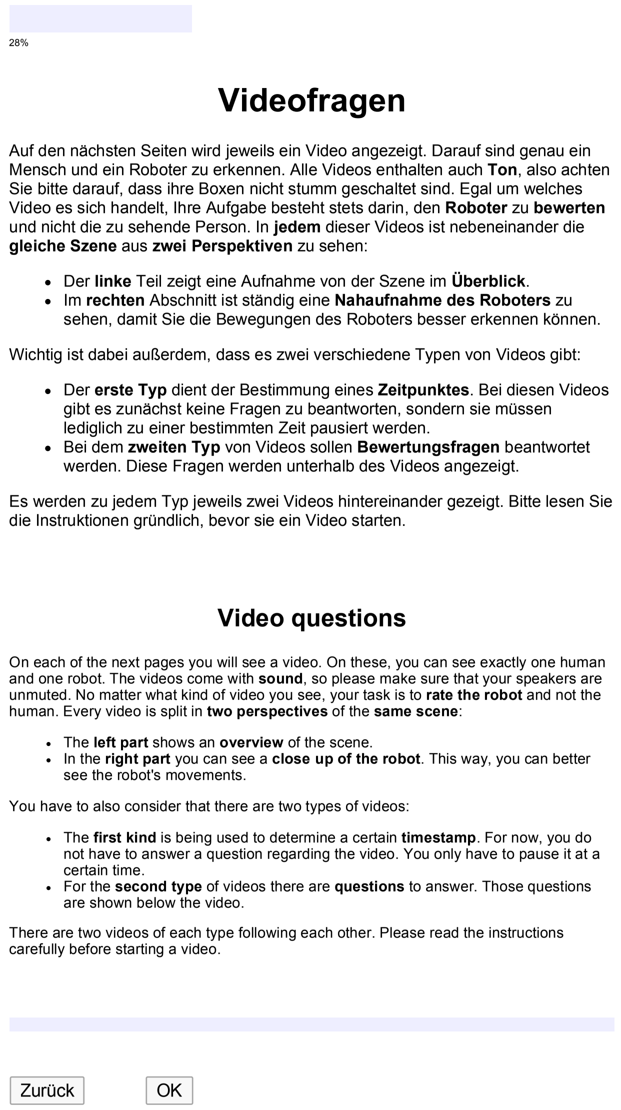
Unter dem Video befinden Sich zwei Knöpfe, mit denen Sie das Video starten und stoppen können. Haben Sie das Video einmal mit dem "Start" Knopf gestartet, können Sie es mit dem "Stop" Knopf wieder anhalten. *Sie haben sonst keine Möglichkeit, das Video zu pausieren, neuzustarten, oder zu wiederholen.*

Below, you can see a video. Your task is to mark a specific time in this video. Please stop the video as soon as you think the following happens: **The robot recognizes that the human wants to communicate to it**. The point the video has been stopped at now marks the time that the robot discovered the human's intention (contacting, communication).

Below the video there are two buttons that you can use to start and stop the video. Once you have started the video with the "Start" button, it can be stopped again using the "Stop" button. *There is no other way to pause, restart or repeat the video at any other time.*

Bitte starten Sie das Video erst **nachdem** Sie die obigen Anweisungen komplett gelesen haben.

Please start the video only **after** you have read the above instructions above.



Videofragen

Auf den nächsten Seiten wird jeweils ein Video angezeigt. Darauf sind genau ein Mensch und ein Roboter zu erkennen. Alle Videos enthalten auch **Ton**, also achten Sie bitte darauf, dass ihre Boxen nicht stumm geschaltet sind. Egal um welches Video es sich handelt, Ihre Aufgabe besteht stets darin, den **Roboter** zu **bewerten** und nicht die zu sehende Person. In **jedem** dieser Videos ist nebeneinander die **gleiche Szene** aus **zwei Perspektiven** zu sehen:

- Der **linke** Teil zeigt eine Aufnahme von der Szene im **Überblick**.
- Im **rechten** Abschnitt ist ständig eine **Nahaufnahme des Roboters** zu sehen, damit Sie die Bewegungen des Roboters besser erkennen können.

Wichtig ist dabei außerdem, dass es zwei verschiedene Typen von Videos gibt:

- Der **erste Typ** dient der Bestimmung eines **Zeitpunktes**. Bei diesen Videos gibt es zunächst keine Fragen zu beantworten, sondern sie müssen lediglich zu einer bestimmten Zeit pausiert werden.
- Bei dem **zweiten Typ** von Videos sollen **Bewertungsfragen** beantwortet werden. Diese Fragen werden unterhalb des Videos angezeigt.

Es werden zu jedem Typ jeweils zwei Videos hintereinander gezeigt. Bitte lesen Sie die Instruktionen gründlich, bevor sie ein Video starten.

Video questions

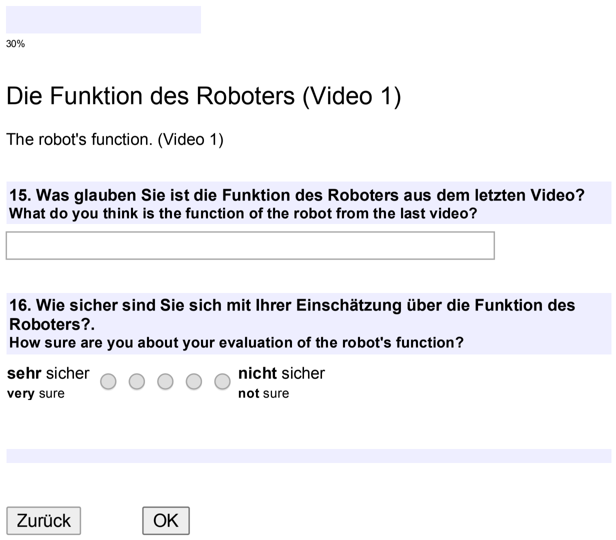
On each of the next pages you will see a video. On these, you can see exactly one human and one robot. The videos come with **sound**, so please make sure that your speakers are unmuted. No matter what kind of video you see, your task is to **rate the robot** and not the human. Every video is split in **two perspectives** of the **same scene**:

- The **left part** shows an **overview** of the scene.
- In the **right part** you can see a **close up of the robot**. This way, you can better see the robot's movements.

You have to also consider that there are two types of videos:

- The **first kind** is being used to determine a certain **timestamp**. For now, you do not have to answer a question regarding the video. You only have to pause it at a certain time.
- For the **second type** of videos there are **questions** to answer. Those questions are shown below the video.

There are two videos of each type following each other. Please read the instructions carefully before starting a video.



30%

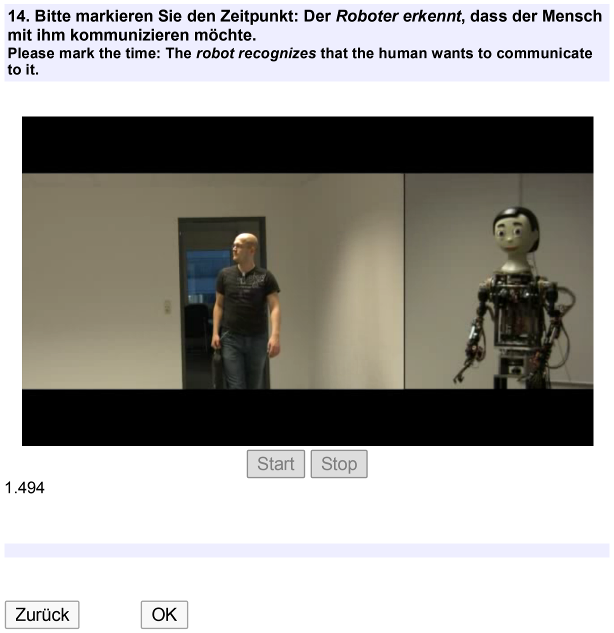
Die Funktion des Roboters (Video 1)

The robot's function. (Video 1)

15. Was glauben Sie ist die Funktion des Roboters aus dem letzten Video?
What do you think is the function of the robot from the last video?

16. Wie sicher sind Sie sich mit Ihrer Einschätzung über die Funktion des Roboters?.
How sure are you about your evaluation of the robot's function?

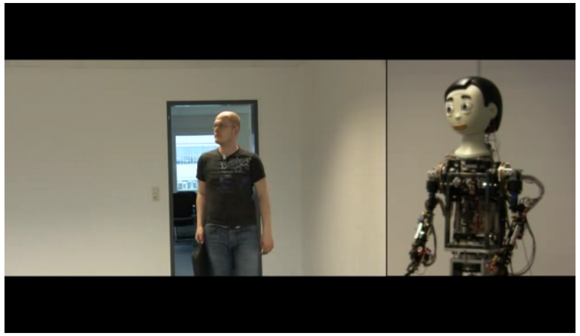
sehr sicher ○ ○ ○ ○ ○ **nicht** sicher
very sure **not** sure



14. Bitte markieren Sie den Zeitpunkt: Der *Roboter erkennt*, dass der Mensch mit ihm kommunizieren möchte.
Please mark the time: The *robot recognizes* that the human wants to communicate to it.

1.494

17. Bitte markieren Sie den Zeitpunkt: Der *Roboter erkennt*, dass der Mensch mit ihm kommunizieren möchte.
Please mark the time: The *robot recognizes* that the human wants to communicate to it.



Start Stop

3.883

Zurück OK

34%

Video 2

Unten auf der Seite sehen Sie ein Video. Ihre Aufgabe besteht darin, einen bestimmten Zeitpunkt in diesem Video zu markieren. Sie stoppen also das Video genau dann, wenn Sie glauben: **Der Roboter erkennt, dass der Mensch mit ihm kommunizieren möchte**. Die Stelle, an der Sie das Video angehalten haben markiert also jetzt den Zeitpunkt, an dem der Roboter die Absicht des Menschen (sprich: Kontaktaufnahme, Kommunikation) erkannt hat.

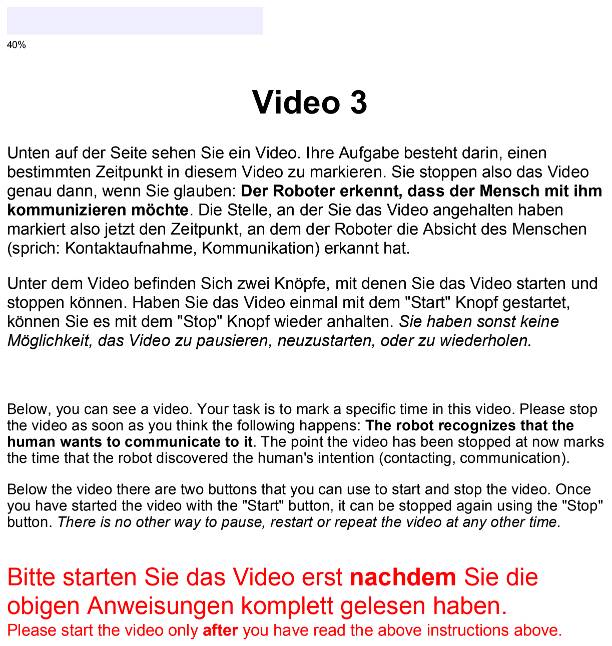
Unter dem Video befinden Sich zwei Knöpfe, mit denen Sie das Video starten und stoppen können. Haben Sie das Video einmal mit dem "Start" Knopf gestartet, können Sie es mit dem "Stop" Knopf wieder anhalten. *Sie haben sonst keine Möglichkeit, das Video zu pausieren, neuzustarten, oder zu wiederholen.*

Below, you can see a video. Your task is to mark a specific time in this video. Please stop the video as soon as you think the following happens: **The robot recognizes that the human wants to communicate to it**. The point the video has been stopped at now marks the time that the robot discovered the human's intention (contacting, communication).

Below the video there are two buttons that you can use to start and stop the video. Once you have started the video with the "Start" button, it can be stopped again using the "Stop" button. *There is no other way to pause, restart or repeat the video at any other time.*

Bitte starten Sie das Video erst **nachdem** Sie die obigen Anweisungen komplett gelesen haben.

Please start the video only **after** you have read the above instructions above.



40%

Video 3

Unten auf der Seite sehen Sie ein Video. Ihre Aufgabe besteht darin, einen bestimmten Zeitpunkt in diesem Video zu markieren. Sie stoppen also das Video genau dann, wenn Sie glauben: **Der Roboter erkennt, dass der Mensch mit ihm kommunizieren möchte**. Die Stelle, an der Sie das Video angehalten haben markiert also jetzt den Zeitpunkt, an dem der Roboter die Absicht des Menschen (sprich: Kontaktaufnahme, Kommunikation) erkannt hat.

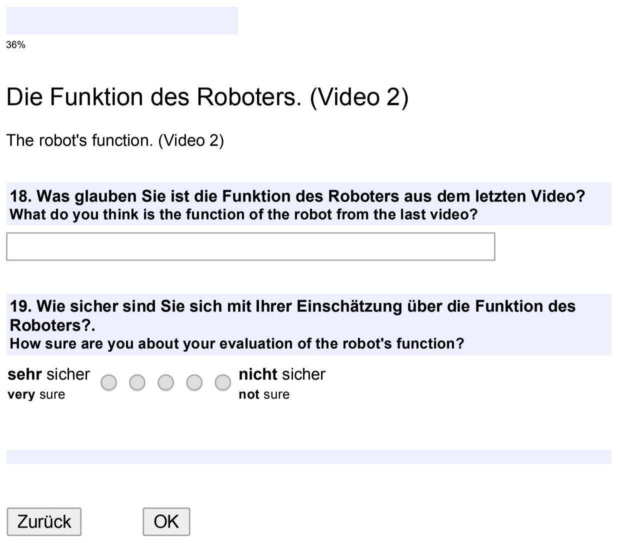
Unter dem Video befinden Sich zwei Knöpfe, mit denen Sie das Video starten und stoppen können. Haben Sie das Video einmal mit dem "Start" Knopf gestartet, können Sie es mit dem "Stop" Knopf wieder anhalten. *Sie haben sonst keine Möglichkeit, das Video zu pausieren, neuzustarten, oder zu wiederholen.*

Below, you can see a video. Your task is to mark a specific time in this video. Please stop the video as soon as you think the following happens: **The robot recognizes that the human wants to communicate to it**. The point the video has been stopped at now marks the time that the robot discovered the human's intention (contacting, communication).

Below the video there are two buttons that you can use to start and stop the video. Once you have started the video with the "Start" button, it can be stopped again using the "Stop" button. *There is no other way to pause, restart or repeat the video at any other time.*

Bitte starten Sie das Video erst **nachdem** Sie die obigen Anweisungen komplett gelesen haben.

Please start the video only **after** you have read the above instructions above.



36%

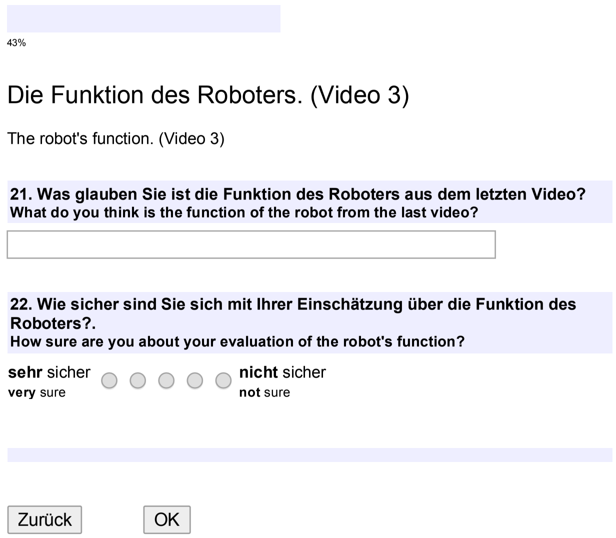
Die Funktion des Roboters. (Video 2)

The robot's function. (Video 2)

18. Was glauben Sie ist die Funktion des Roboters aus dem letzten Video?
What do you think is the function of the robot from the last video?

19. Wie sicher sind Sie sich mit Ihrer Einschätzung über die Funktion des Roboters?.
How sure are you about your evaluation of the robot's function?

sehr sicher ○ ○ ○ ○ ○ **nicht** sicher
very sure **not** sure



43%

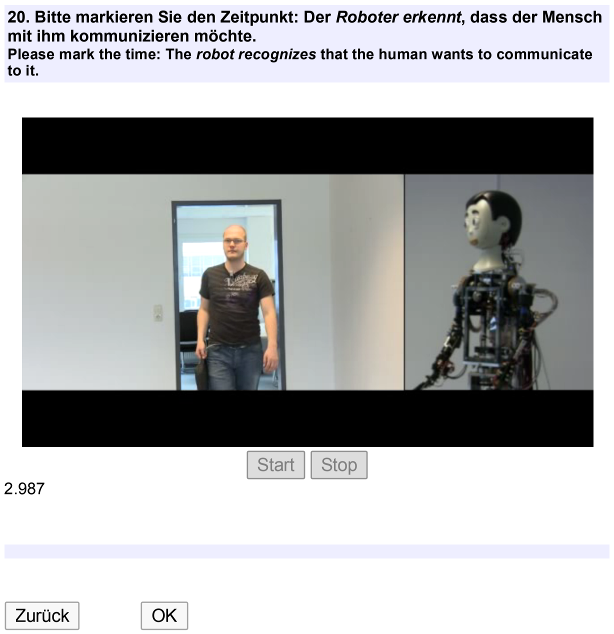
Die Funktion des Roboters. (Video 3)

The robot's function. (Video 3)

21. Was glauben Sie ist die Funktion des Roboters aus dem letzten Video?
What do you think is the function of the robot from the last video?

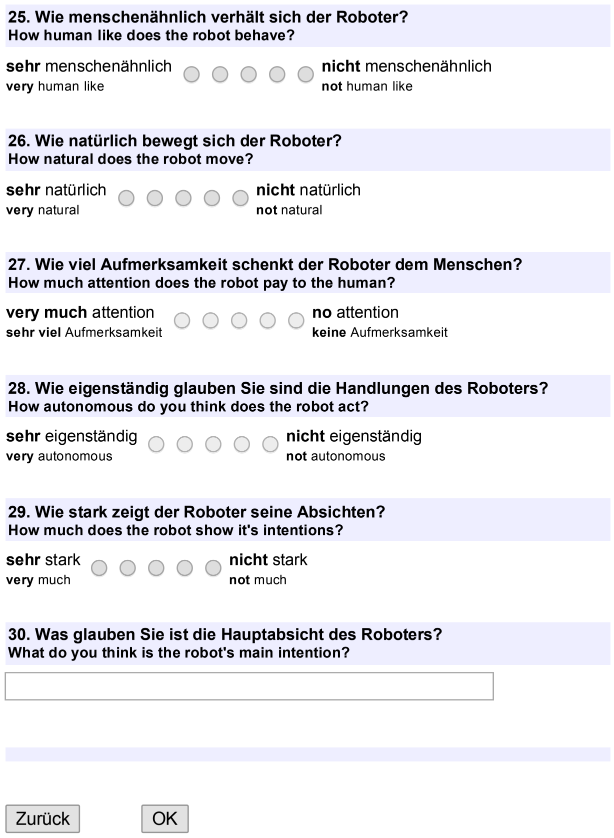
22. Wie sicher sind Sie sich mit Ihrer Einschätzung über die Funktion des Roboters?.
How sure are you about your evaluation of the robot's function?

sehr sicher ○ ○ ○ ○ ○ **nicht** sicher
very sure **not** sure



20. Bitte markieren Sie den Zeitpunkt: Der *Roboter erkennt*, dass der Mensch mit ihm kommunizieren möchte.
Please mark the time: The *robot recognizes* that the human wants to communicate to it.

2.987



25. Wie menschenähnlich verhält sich der Roboter?
How human like does the robot behave?

sehr menschenähnlich ○ ○ ○ ○ ○ **nicht** menschenähnlich
very human like — **not** human like

26. Wie natürlich bewegt sich der Roboter?
How natural does the robot move?

sehr natürlich ○ ○ ○ ○ ○ **nicht** natürlich
very natural — **not** natural

27. Wie viel Aufmerksamkeit schenkt der Roboter dem Menschen?
How much attention does the robot pay to the human?

very much attention ○ ○ ○ ○ ○ **no** attention
sehr viel Aufmerksamkeit — **keine** Aufmerksamkeit

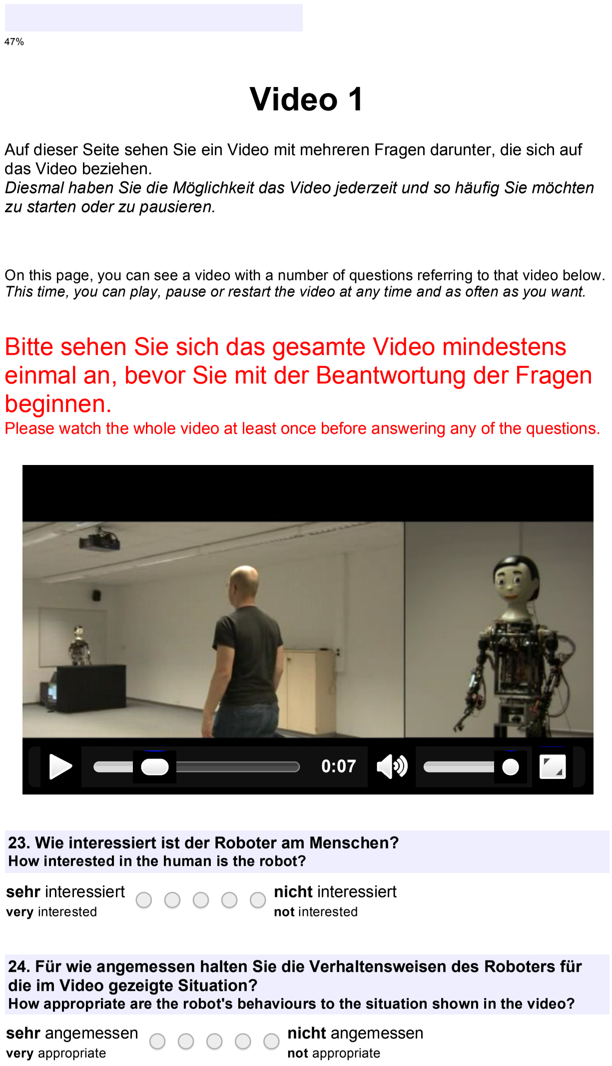
28. Wie eigenständig glauben Sie sind die Handlungen des Roboters?
How autonomous do you think does the robot act?

sehr eigenständig ○ ○ ○ ○ ○ **nicht** eigenständig
very autonomous — **not** autonomous

29. Wie stark zeigt der Roboter seine Absichten?
How much does the robot show it's intentions?

sehr stark ○ ○ ○ ○ ○ **nicht** stark
very much — **not** much

30. Was glauben Sie ist die Hauptabsicht des Roboters?
What do you think is the robot's main intention?



47%

Video 1

Auf dieser Seite sehen Sie ein Video mit mehreren Fragen darunter, die sich auf das Video beziehen.
Diesmal haben Sie die Möglichkeit das Video jederzeit und so häufig Sie möchten zu starten oder zu pausieren.

On this page, you can see a video with a number of questions referring to that video below.
This time, you can play, pause or restart the video at any time and as often as you want.

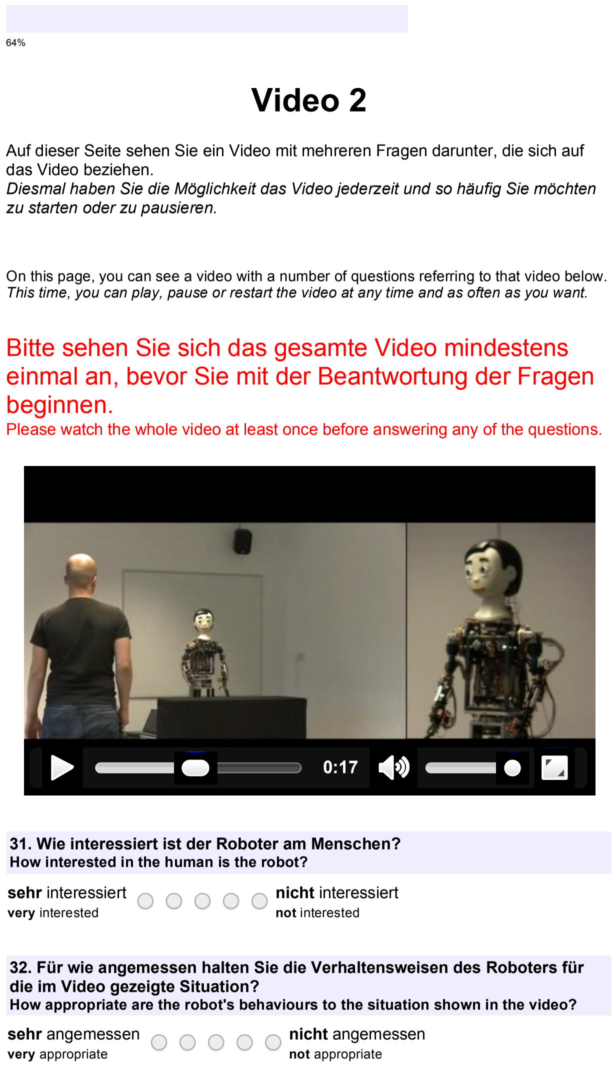
Bitte sehen Sie sich das gesamte Video mindestens einmal an, bevor Sie mit der Beantwortung der Fragen beginnen.
Please watch the whole video at least once before answering any of the questions.

23. Wie interessiert ist der Roboter am Menschen?
How interested in the human is the robot?

sehr interessiert ○ ○ ○ ○ ○ **nicht** interessiert
very interested — **not** interested

24. Für wie angemessen halten Sie die Verhaltensweisen des Roboters für die im Video gezeigte Situation?
How appropriate are the robot's behaviours to the situation shown in the video?

sehr angemessen ○ ○ ○ ○ ○ **nicht** angemessen
very appropriate — **not** appropriate



64%

Video 2

Auf dieser Seite sehen Sie ein Video mit mehreren Fragen darunter, die sich auf das Video beziehen.
Diesmal haben Sie die Möglichkeit das Video jederzeit und so häufig Sie möchten zu starten oder zu pausieren.

On this page, you can see a video with a number of questions referring to that video below.
This time, you can play, pause or restart the video at any time and as often as you want.

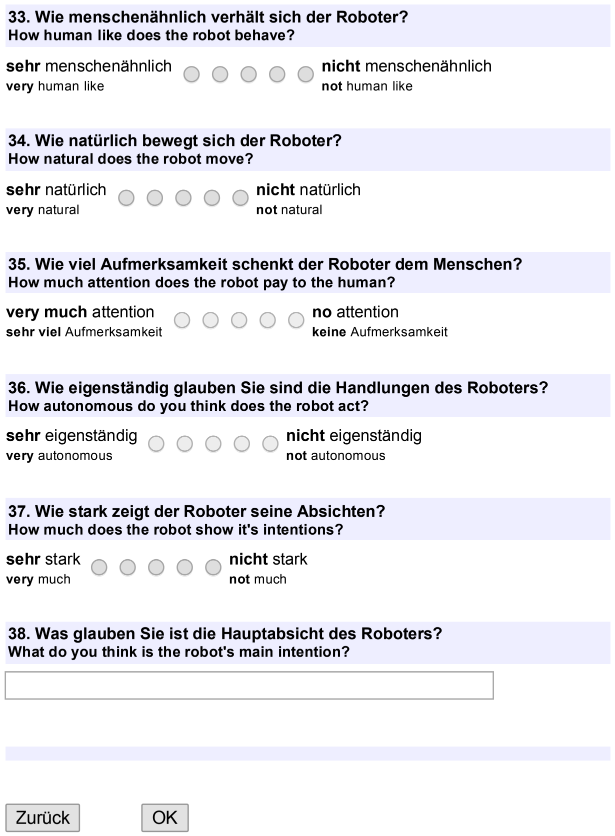
Bitte sehen Sie sich das gesamte Video mindestens einmal an, bevor Sie mit der Beantwortung der Fragen beginnen.
Please watch the whole video at least once before answering any of the questions.

31. Wie interessiert ist der Roboter am Menschen?
How interested in the human is the robot?

sehr interessiert ○ ○ ○ ○ ○ **nicht** interessiert
very interested **not** interested

32. Für wie angemessen halten Sie die Verhaltensweisen des Roboters für die im Video gezeigte Situation?
How appropriate are the robot's behaviours to the situation shown in the video?

sehr angemessen ○ ○ ○ ○ ○ **nicht** angemessen
very appropriate **not** appropriate



33. Wie menschenähnlich verhält sich der Roboter?
How human like does the robot behave?

sehr menschenähnlich ○ ○ ○ ○ ○ **nicht** menschenähnlich
very human like **not** human like

34. Wie natürlich bewegt sich der Roboter?
How natural does the robot move?

sehr natürlich ○ ○ ○ ○ ○ **nicht** natürlich
very natural **not** natural

35. Wie viel Aufmerksamkeit schenkt der Roboter dem Menschen?
How much attention does the robot pay to the human?

very much attention ○ ○ ○ ○ ○ **no** attention
sehr viel Aufmerksamkeit **keine** Aufmerksamkeit

36. Wie eigenständig glauben Sie sind die Handlungen des Roboters?
How autonomous do you think does the robot act?

sehr eigenständig ○ ○ ○ ○ ○ **nicht** eigenständig
very autonomous **not** autonomous

37. Wie stark zeigt der Roboter seine Absichten?
How much does the robot show it's intentions?

sehr stark ○ ○ ○ ○ ○ **nicht** stark
very much **not** much

38. Was glauben Sie ist die Hauptabsicht des Roboters?
What do you think is the robot's main intention?

81%

Video 3

Auf dieser Seite sehen Sie ein Video mit mehreren Fragen darunter, die sich auf das Video beziehen.
Diesmal haben Sie die Möglichkeit das Video jederzeit und so häufig Sie möchten zu starten oder zu pausieren.

On this page, you can see a video with a number of questions referring to that video below.
This time, you can play, pause or restart the video at any time and as often as you want.

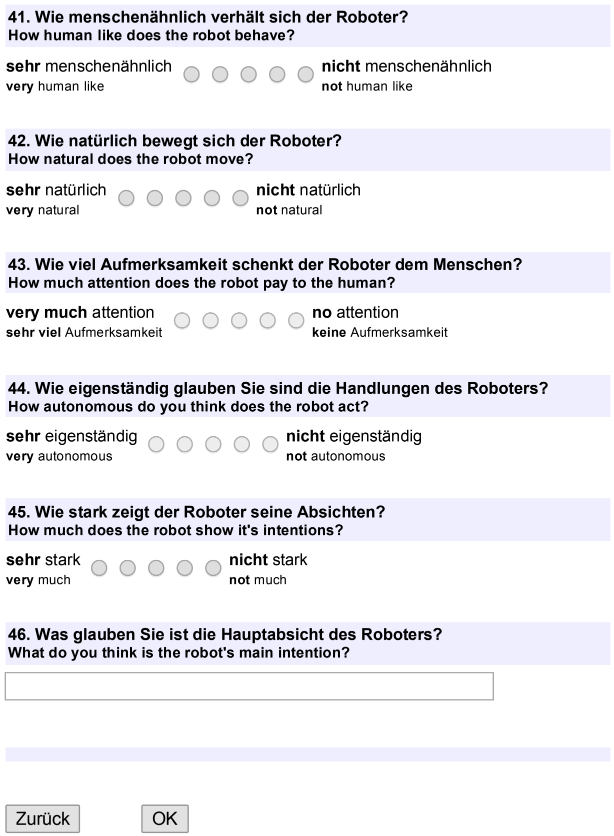
Bitte sehen Sie sich das gesamte Video mindestens einmal an, bevor Sie mit der Beantwortung der Fragen beginnen.
Please watch the whole video at least once before answering any of the questions.

39. Wie interessiert ist der Roboter am Menschen?
How interested in the human is the robot?

sehr interessiert ○ ○ ○ ○ ○ **nicht** interessiert
very interested — **not** interested

40. Für wie angemessen halten Sie die Verhaltensweisen des Roboters für die im Video gezeigte Situation?
How appropriate are the robot's behaviours to the situation shown in the video?

sehr angemessen ○ ○ ○ ○ ○ **nicht** angemessen
very appropriate — **not** appropriate



41. Wie menschenähnlich verhält sich der Roboter?
How human like does the robot behave?

sehr menschenähnlich ○ ○ ○ ○ ○ **nicht** menschenähnlich
very human like — **not** human like

42. Wie natürlich bewegt sich der Roboter?
How natural does the robot move?

sehr natürlich ○ ○ ○ ○ ○ **nicht** natürlich
very natural — **not** natural

43. Wie viel Aufmerksamkeit schenkt der Roboter dem Menschen?
How much attention does the robot pay to the human?

very much attention ○ ○ ○ ○ ○ **no** attention
sehr viel Aufmerksamkeit — **keine** Aufmerksamkeit

44. Wie eigenständig glauben Sie sind die Handlungen des Roboters?
How autonomous do you think does the robot act?

sehr eigenständig ○ ○ ○ ○ ○ **nicht** eigenständig
very autonomous — **not** autonomous

45. Wie stark zeigt der Roboter seine Absichten?
How much does the robot show it's intentions?

sehr stark ○ ○ ○ ○ ○ **nicht** stark
very much — **not** much

46. Was glauben Sie ist die Hauptabsicht des Roboters?
What do you think is the robot's main intention?

98%

Ende

Falls Sie noch Anmerkungen, Kommentare oder Kritik zu dieser Studie haben, dürfen Sie diese jetzt in dem unten stehenden Feld anbringen.

Finish

If you have any comments or criticism about this study, then please tell us using the field below.

47. Kommentare:
Comments:

Zurück Abschicken

F. Questionnaire Live Interaction

Willkommen zur Roboter Studie mit iCub!

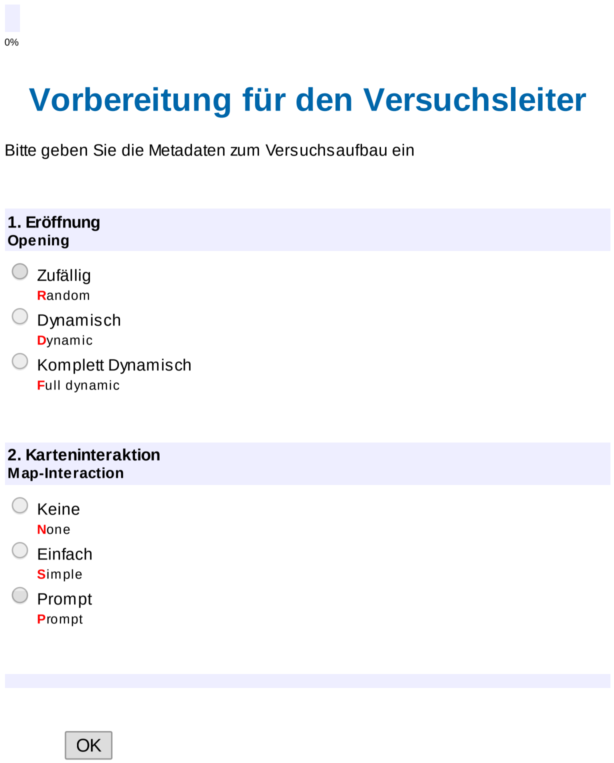
Im Folgenden haben Sie die Möglichkeit bei einer **Studie** zur **Mensch-Roboter Interaktion** teilzunehmen. Diese Studie gliedert sich in **zwei Teile**: Zunächst bitten wir Sie, sich mit dem Roboter iCub zu unterhalten, der sich im Nebenraum befindet. Dabei besteht Ihre Aufgabe darin, Informationen vom Roboter zu erhalten. Im zweiten Teil sollen Sie dann Fragen an diesem Computer beantworten, welche sich auf die erlebte Situation beziehen.

Die Gesamtdauer der Studie wird nicht länger als 20 Minuten betragen.

Welcome to the robot study with iCub!

Today, you will have the possibility to contribute to a **study** on **human-robot interaction**. The study is divided in **two parts**: Firstly, we ask you to talk to the iCub robot which you can find in the next room. Your task will consist of retrieving information from the robot. Secondly, you should answer questions here on the PC which are related to the just experienced situation.

The overal duration of the study will not exceed 20 minutes.



Vorbereitung für den Versuchsleiter

Bitte geben Sie die Metadaten zum Versuchsaufbau ein

1. Eröffnung
Opening

○ Zufällig
Random

○ Dynamisch
Dynamic

○ Komplett Dynamisch
Full dynamic

2. Karteninteraktion
Map-Interaction

○ Keine
None

○ Einfach
Simple

○ Prompt
Prompt

5%

[english below]

Einverständniserklärung

Hiermit erkläre ich mich damit einverstanden, dass im Rahmen dieser Studie zum Verhältnis zwischen Mensch und Roboter u.a. **personenbezogene Daten** inklusive Bild- und Tonmaterial von mir erfasst werden. Diese Daten werden **vertraulich** und ausschließlich zur **wissenschaftlichen Auswertung** der Studie verwendet. **Anonymisierte Ergebnisse** der Studie können auf wissenschaftlichen Veranstaltungen wie Konferenzen oder Kolloquien gezeigt werden. Mir ist bewusst, dass meine Antworten mit personenbezogenen Daten (z.B. Alter, Geschlecht, Studiengang oder Beruf) in Verbindung gebracht werden können.

Die Einwilligung über die Verwendung dieser Angaben ist jederzeit widerruflich. Im Falle des Widerrufs dürfen personenbezogene Daten zukünftig nicht mehr für die oben genannten Zwecke verwendet werden und sind unverzüglich zu löschen. Soweit die Einwilligung nicht widerrufen wird, gilt sie zeitlich unbeschränkt. Die Einwilligung ist freiwillig. Aus der Verweigerung der Einwilligung oder ihrem Widerruf entstehen keine Nachteile.

Privacy statement

I hereby agree that among others **personal data** including video and audio recordings is collected during this study on the relationship between humans and robots. This data is handled **confidentially** and will only be used for the **scientific analysis** of this study. **Anonymous results** may be shown on scientific events like conferences or colloqiums. I am aware that my answers can be related to personal data of mine (e.g. age, sex, field of study).

This agreement can be revoked at any time. In this case, personal data may not be used anymore for the above purposes - they have to be deleted immediately. If the agreement is not revoked, it is valid infinitely. This agreement is voluntarily. There are no disadvantages caused by the agreement or its revokation.

☐ Ich bin ausreichend über den Datenschutz informiert worden und bin damit einverstanden, dass meine Daten zu wissenschaftlichen Zwecken genutzt werden dürfen.

I have been informed well about the privacy policy and I agree to the scientific usage of my data.

Zurück OK

5%

[english below]

Teil 1 (Einführung)

Bitte lesen Sie diese Seite besonders aufmerksam.

Ihre Aufgabe

Im Raum nebenan befindet sich der Roboter iCub, der die Funktion eines Rezeptionisten ausübt und Auskunft über Räume in diesem Gebäude geben kann. Stellen Sie sich vor, sie suchen den Weg zu einem dieser Räume. Ihre Aufgabe besteht nun darin, den Roboter zu fragen, wie Sie die Wege zu den Räumen finden können.

Bitte fragen Sie iCub **nacheinander** nach diesen Orten:

1. Auditorium (english: auditorium)
2. Zentraler Innenhof
3. Äußerer Innenhof

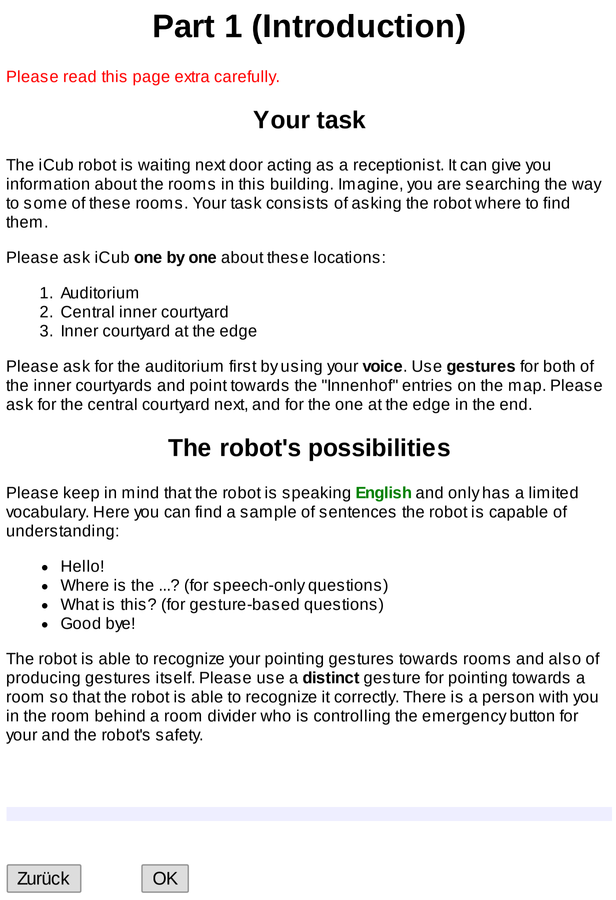
Bitte fragen Sie zuerst nach dem Auditorium, indem Sie den Roboter direkt nach diesem Raum **fragen** ohne auf die Karte zu zeigen. Für die beiden Innenhöfe benutzen Sie bitte **Gesten**, zeigen auf den jeweiligen Abschnitt auf der Karte und fragen "What is this?". Bitte fragen Sie dann nach dem Innenhof in der Mitte des Gebäudes und zum Schluß nach dem Innenhof, der etwas am Rand liegt.

Die Möglichkeiten des Roboters

Beachten Sie bitte, dass der Roboter nur **englisch** versteht und nur über einen beschränkten Wortschatz verfügt. Sprechen Sie deshalb bitte nur englisch mit iCub. Im Folgenden sehen Sie eine Auswahl an Sätzen, die der Roboter versteht:

- Hello!
- Where is the ...? (Für Fragen per Sprache)
- What is this? (Für Fragen per Geste)
- Good bye!

Der Roboter ist in der Lage, Ihre Zeigegesten auf Räume zu deuten und auch selbstständig Gesten zu produzieren. Bitte zeigen Sie **deutlich** auf den Raum, damit der Roboter Ihre Geste richtig erkennen kann. Zur Ihrer Sicherheit und auch der des Roboters sitzt eine Person, die den Notschalter des Roboters bedient, mit Ihnen im Raum hinter einer Trennwand.



Part 1 (Introduction)

Please read this page extra carefully.

Your task

The iCub robot is waiting next door acting as a receptionist. It can give you information about the rooms in this building. Imagine, you are searching the way to some of these rooms. Your task consists of asking the robot where to find them.

Please ask iCub **one by one** about these locations:

1. Auditorium
2. Central inner courtyard
3. Inner courtyard at the edge

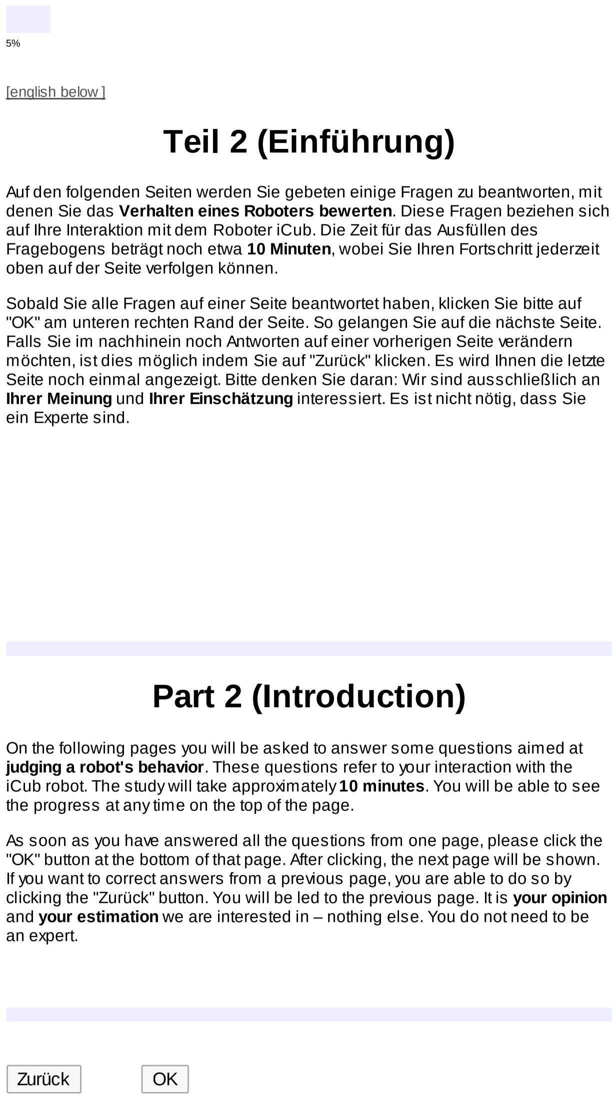
Please ask for the auditorium first by using your **voice**. Use **gestures** for both of the inner courtyards and point towards the "Innenhof" entries on the map. Please ask for the central courtyard next, and for the one at the edge in the end.

The robot's possibilities

Please keep in mind that the robot is speaking **English** and only has a limited vocabulary. Here you can find a sample of sentences the robot is capable of understanding:

- Hello!
- Where is the ...? (for speech-only questions)
- What is this? (for gesture-based questions)
- Good bye!

The robot is able to recognize your pointing gestures towards rooms and also of producing gestures itself. Please use a **distinct** gesture for pointing towards a room so that the robot is able to recognize it correctly. There is a person with you in the room behind a room divider who is controlling the emergency button for your and the robot's safety.



Teil 2 (Einführung)

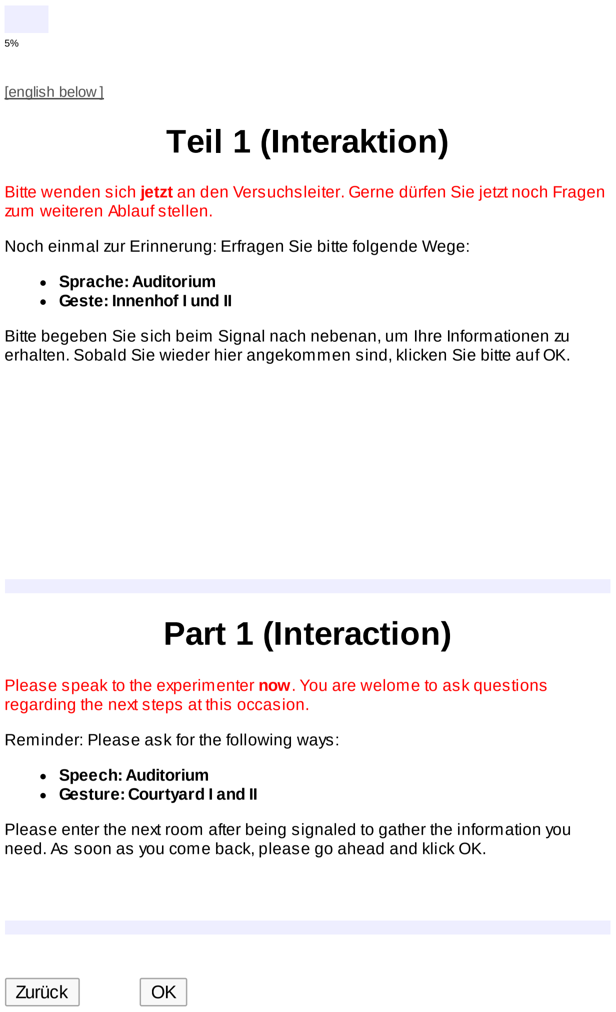
Auf den folgenden Seiten werden Sie gebeten einige Fragen zu beantworten, mit denen Sie das **Verhalten eines Roboters bewerten**. Diese Fragen beziehen sich auf Ihre Interaktion mit dem Roboter iCub. Die Zeit für das Ausfüllen des Fragebogens beträgt noch etwa **10 Minuten**, wobei Sie Ihren Fortschritt jederzeit oben auf der Seite verfolgen können.

Sobald Sie alle Fragen auf einer Seite beantwortet haben, klicken Sie bitte auf "OK" am unteren rechten Rand der Seite. So gelangen Sie auf die nächste Seite. Falls Sie im nachhinein noch Antworten auf einer vorherigen Seite verändern möchten, ist dies möglich indem Sie auf "Zurück" klicken. Es wird Ihnen die letzte Seite noch einmal angezeigt. Bitte denken Sie daran: Wir sind ausschließlich an **Ihrer Meinung** und **Ihrer Einschätzung** interessiert. Es ist nicht nötig, dass Sie ein Experte sind.

Part 2 (Introduction)

On the following pages you will be asked to answer some questions aimed at **judging a robot's behavior**. These questions refer to your interaction with the iCub robot. The study will take approximately **10 minutes**. You will be able to see the progress at any time on the top of the page.

As soon as you have answered all the questions from one page, please click the "OK" button at the bottom of that page. After clicking, the next page will be shown. If you want to correct answers from a previous page, you are able to do so by clicking the "Zurück" button. You will be led to the previous page. It is **your opinion** and **your estimation** we are interested in – nothing else. You do not need to be an expert.



Teil 1 (Interaktion)

Bitte wenden sich **jetzt** an den Versuchsleiter. Gerne dürfen Sie jetzt noch Fragen zum weiteren Ablauf stellen.

Noch einmal zur Erinnerung: Erfragen Sie bitte folgende Wege:

- **Sprache: Auditorium**
- **Geste: Innenhof I und II**

Bitte begeben Sie sich beim Signal nach nebenan, um Ihre Informationen zu erhalten. Sobald Sie wieder hier angekommen sind, klicken Sie bitte auf OK.

Part 1 (Interaction)

Please speak to the experimenter **now**. You are welome to ask questions regarding the next steps at this occasion.

Reminder: Please ask for the following ways:

- **Speech: Auditorium**
- **Gesture: Courtyard I and II**

Please enter the next room after being signaled to gather the information you need. As soon as you come back, please go ahead and klick OK.

11%

Allgemeine Fragen

Auf der dieser Seite finden Sie allgemeine Fragen zu ihrer Begegnung mit dem Roboter iCub. Lassen Sie bitte Ihren **Gesamteindruck** vom Betreten des Raumes bis zum Verlassen in Ihre Bewertung mit einfließen.

General Questions

On this page, you can find general questions about the encounter with the iCub robot. Please rate the **complete situation** from entering the room until you came back here.

6. Wie interessiert ist der Roboter an Ihnen?
How interested in you is the robot?

sehr interessiert **very** interested	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** interessiert **not** interested

7. Für wie angemessen halten Sie die Verhaltensweisen des Roboters für die gerade erlebte Situation?
How appropriate to the experienced situation are the robot's behaviours?

sehr angemessen **very** appropriate	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** angemessen **not** appropriate

8. Wie menschenähnlich verhält sich der Roboter?
How human-like does the robot behave?

sehr menschenähnlich **very** human-like	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** menschenähnlich **not** human-like

9. Wie natürlich bewegt sich der Roboter?
How natural does the robot move?

sehr natürlich **very** natural	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** natürlich **not** natural

10. Wie viel Aufmerksamkeit schenkt der Roboter Ihnen?
How much attention does the robot pay to you?

very much attention **sehr viel** Aufmerksamkeit	○ 4	○ 3	○ 2	○ 1	○ 0	**no** attention **keine** Aufmerksamkeit

---

5%

Erfahrung mit Computern und Robotern

Experience with robots and computers

3. Bitte bewerten Sie ihre Erfahrung mit *Computern* im Allgemeinen.
Please rate your general level of experience with *computers*.

sehr erfahren **very** experienced	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** erfahren **not** experienced

4. Bitte nennen Sie uns *alle* Ihnen bekannte Roboter.
Please tell us *every* robot you know.

☐ Asimo
☐ Geminoid
☐ Kismet
☐ Barthoc
☐ Flobi
☐ BIRON
☐ Roboter aus Filmen
Robots from movies
☐ Spielzeugroboter
Toy robots
☐ Industrieroboter
Industrial robots
☐ Andere Roboter, bitte angeben
Other robots, please specify ______

5. Bitte bewerten Sie ihre Erfahrung mit Robotern im Allgemeinen.
Please rate your level of experience with robots.

sehr erfahren **very** experienced	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** erfahren **not** experienced

Zurück OK

30%

Fragen zur Kontaktaufnahme

Auf dieser Seite finden Sie Fragen, die sich mit dem **Einstieg** in die Interaktion mit dem Roboter beschäftigen.

Questions about making contact

On this page, you can find questions regarding the **opening** of your interaction with the robot.

14. Wann glauben Sie, bemerkte der Roboter, dass Sie mit ihm interagieren möchten?
When do you think the robot noticed that you wanted to interact with it?

- ○ Sofort beim Betreten des Raumes
  Directly, as I entered the room
- ○ Während Ich mich dem Roboter genähert habe
  While I was approaching the robot
- ○ Als ich direkt vor dem Tisch stand
  As I stood directly in front of the table
- ○ Bei der Begrüßung
  At the time of greeting
- ○ Gar nicht
  Never
- ○ Andere, bitte eintragen
  Other, please specify ☐

15. Wie stark zeigte der Roboter seine Bereitschaft zur Interaktion während Sie sich ihm näherten?
How much did the robot show it's willingness to interact while you were approaching it?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

16. Wie stark fühlten Sie sich ermuntert mit dem Roboter zu sprechen?
How much did did you feel encouraged to speak to the robot?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

11. Für wie eigenständig halten Sie die Handlungen des Roboters?
How autonomous do you think does the robot act?

sehr eigenständig **very** autonomous	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** eigenständig **not** autonomous

12. Wie stark zeigt der Roboter seine Absichten?
How much does the robot show it's intentions?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

13. Gab es unerwartete Verhaltensweisen des Roboters?
Did you experience any unexpected robot behaviors?

☐

Zurück OK

48%

Fragen zur Interaktion am Tisch

Auf der dieser Seite finden Sie Fragen, die sich mit der **direkten Interkation am Tisch** zwischen Ihnen und dem Roboter beschäftigen.

Questions about the interaction at the table

On this page, you can find questions regarding the **direct interaction at the table** between you and the robot.

22. Als wie hilfreich würden Sie die Auskünfte des Roboters insgesamt bezeichnen?
How helpful would you rate the robot's information?

sehr hilfreich **very** helpful	○ 4	○ 3	○ 2	○ 1	○ 0	**nicht** hilfreich **not** helpful

23. Wie stark trug Ihrer Meinung nach die Gestik des Roboters zum Informationsgehalt seiner Auskunft bei?
How much did the robot's gestures support the content of its information?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

24. Wie stark trugen Ihrer Meinung nach die Blicke des Roboters zum Informationsgehalt seiner Auskunft bei?
How much did the robot's gaze support the content of its information?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

25. Wie stark trug Ihrer Meinung nach die Karte auf dem Tisch zum Informationsgehalt der Auskunft bei?
How much did the map on the table support the content of the information?

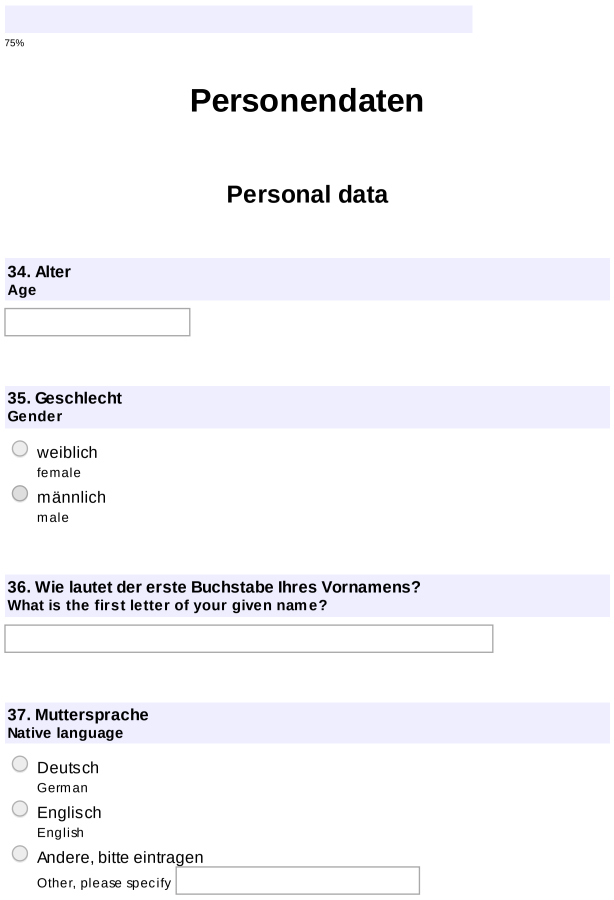
sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

---

17. Hat der Roboter Sie selbstständig begrüßt?
Did the robot greet you autonomously?

- ○ Ja
  Yes
- ○ Nein
  No
- ○ Ich weiß es nicht
  I don't know

Zurück OK



75%

Personendaten

Personal data

34. Alter
Age

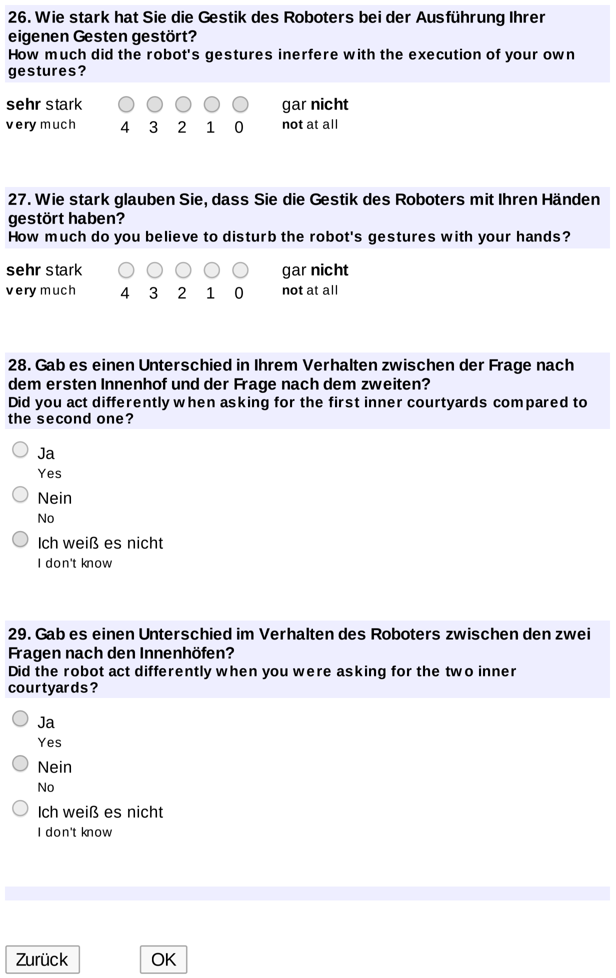
35. Geschlecht
Gender

- ○ weiblich
  female
- ○ männlich
  male

36. Wie lautet der erste Buchstabe Ihres Vornamens?
What is the first letter of your given name?

37. Muttersprache
Native language

- ○ Deutsch
  German
- ○ Englisch
  English
- ○ Andere, bitte eintragen
  Other, please specify



26. Wie stark hat Sie die Gestik des Roboters bei der Ausführung Ihrer eigenen Gesten gestört?
How much did the robot's gestures inerfere with the execution of your own gestures?

sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

27. Wie stark glauben Sie, dass Sie die Gestik des Roboters mit Ihren Händen gestört haben?
How much do you believe to disturb the robot's gestures with your hands?

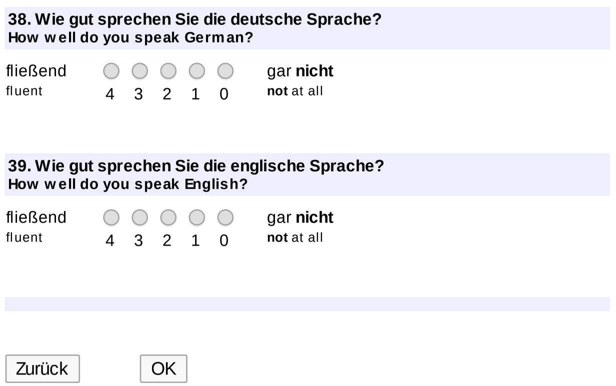
sehr stark **very** much	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

28. Gab es einen Unterschied in Ihrem Verhalten zwischen der Frage nach dem ersten Innenhof und der Frage nach dem zweiten?
Did you act differently when asking for the first inner courtyards compared to the second one?

- ○ Ja
  Yes
- ○ Nein
  No
- ○ Ich weiß es nicht
  I don't know

29. Gab es einen Unterschied im Verhalten des Roboters zwischen den zwei Fragen nach den Innenhöfen?
Did the robot act differently when you were asking for the two inner courtyards?

- ○ Ja
  Yes
- ○ Nein
  No
- ○ Ich weiß es nicht
  I don't know

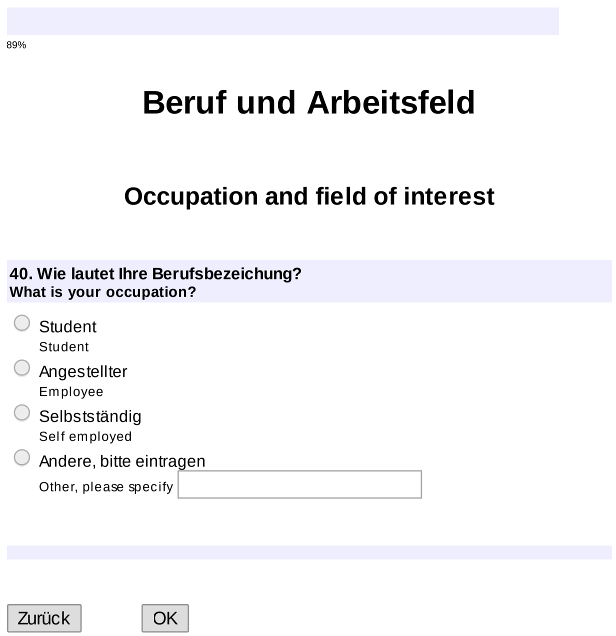


38. Wie gut sprechen Sie die deutsche Sprache?
How well do you speak German?

fließend fluent	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all

39. Wie gut sprechen Sie die englische Sprache?
How well do you speak English?

fließend fluent	○ 4	○ 3	○ 2	○ 1	○ 0	gar **nicht** **not** at all



89%

Beruf und Arbeitsfeld

Occupation and field of interest

40. Wie lautet Ihre Berufsbezeichung?
What is your occupation?

○ Student
Student
○ Angestellter
Employee
○ Selbstständig
Self employed
○ Andere, bitte eintragen
Other, please specify

98%

Ende

Falls Sie noch Anmerkungen, Kommentare oder Kritik zu dieser Studie haben, dürfen Sie diese jetzt in dem unten stehenden Feld anbringen.

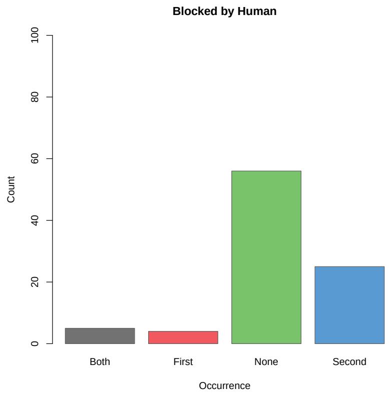
Finish

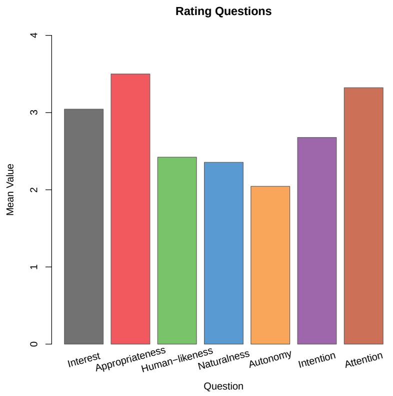
If you have any comments or criticism about this study, then please tell us using the field below .

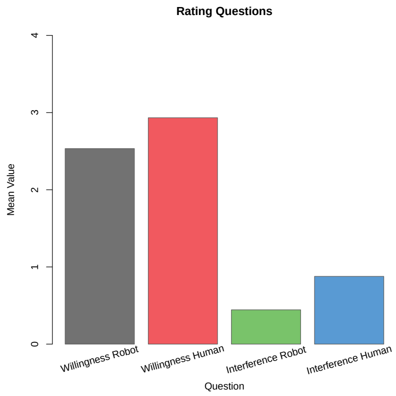
44. Kommentare:
Comments:

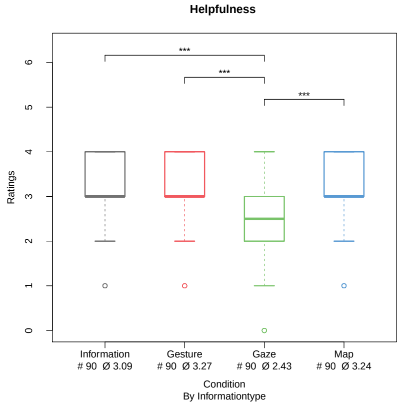
Zurück Abschicken

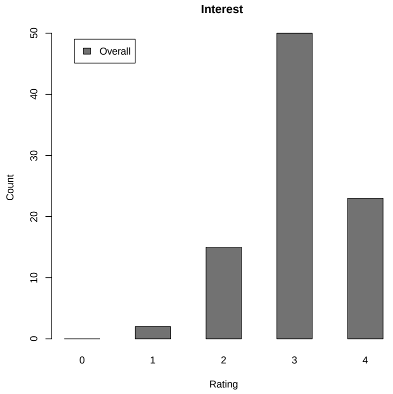
G. Evaluation Live Interaction

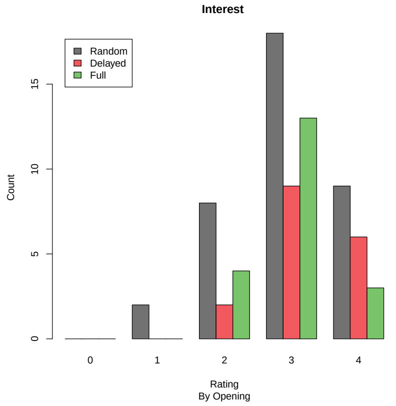

Blocked by Human
Count
0
20
40
60
80
100
Both
First
None
Second
Occurrence

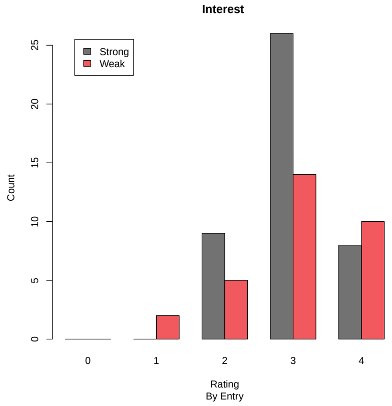

Rating Questions
Mean Value
0
1
2
3
4
Interest
Appropriateness
Human−likeness
Naturalness
Autonomy
Intention
Attention
Question

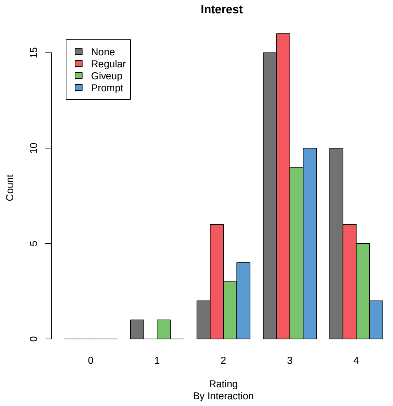

Rating Questions
Mean Value
0
1
2
3
4
Willingness Robot
Willingness Human
Interference Robot
Interference Human
Question

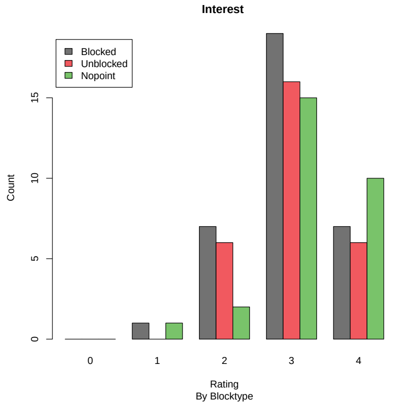

Helpfulness
Ratings
0
1
2
3
4
5
6
***
***
***
Information
90 Ø 3.09
Gesture
90 Ø 3.27
Gaze
90 Ø 2.43
Map
90 Ø 3.24
Condition
By Informationtype

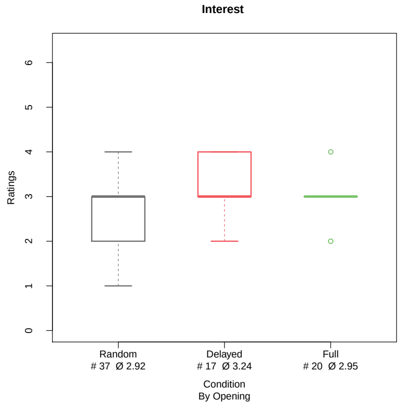

Interest
Count
0
10
20
30
40
50
Overall
0
1
2
3
4
Rating

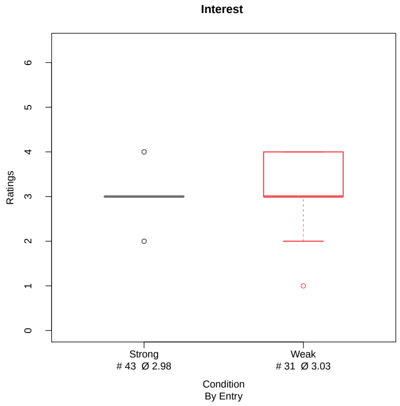

Interest
Count
0
5
10
15
Random
Delayed
Full
0
1
2
3
4
Rating
By Opening

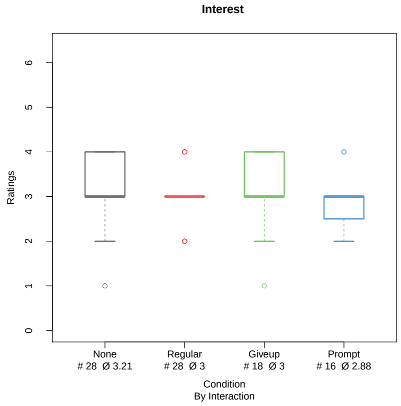

Interest
Count
0
5
10
15
20
25
Strong
Weak
0
1
2
3
4
Rating
By Entry

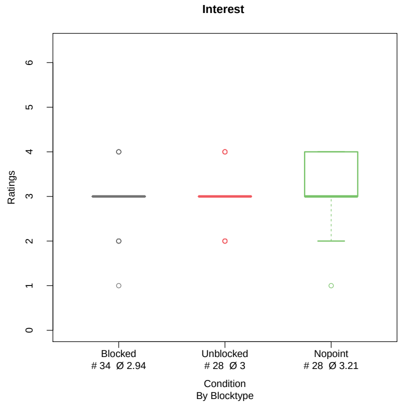

Interest
Count
0
5
10
15
None
Regular
Giveup
Prompt
0
1
2
3
4
Rating
By Interaction

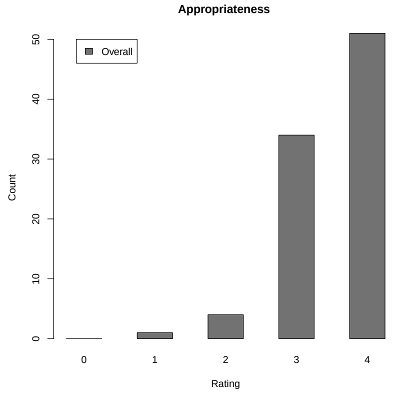

Interest
Count
0
5
10
15
Blocked
Unblocked
Nopoint
0
1
2
3
4
Rating
By Blocktype

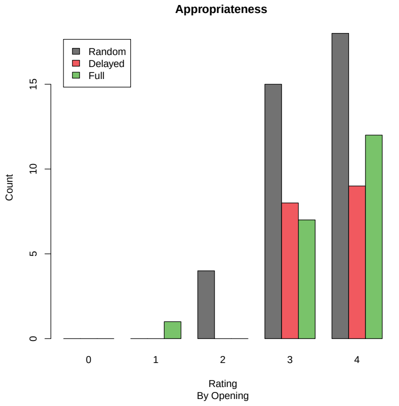

Interest
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.92
Delayed
17 Ø 3.24
Full
20 Ø 2.95
Condition
By Opening

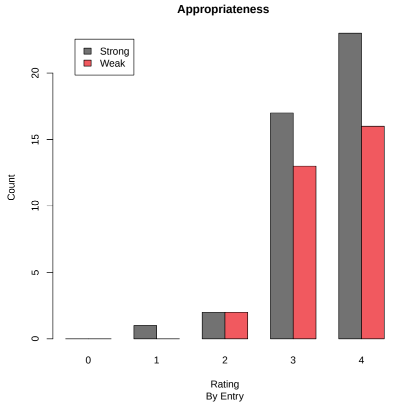

Interest
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.98
Weak
31 Ø 3.03
Condition
By Entry

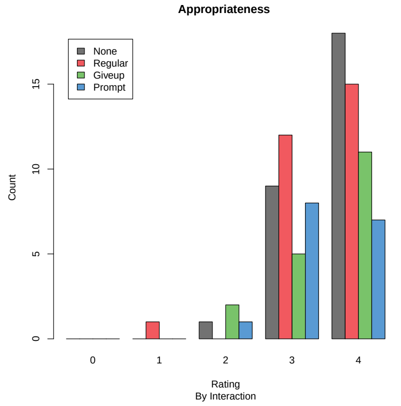

Interest
Ratings
0
1
2
3
4
5
6
None
28 Ø 3.21
Regular
28 Ø 3
Giveup
18 Ø 3
Prompt
16 Ø 2.88
Condition
By Interaction

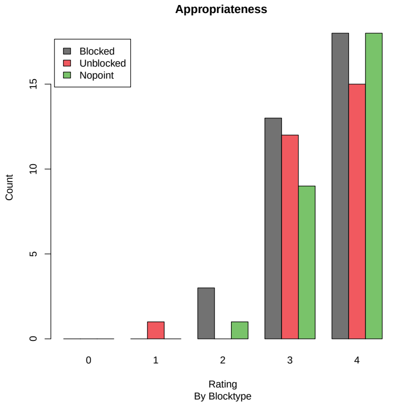

Interest
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.94
Unblocked
28 Ø 3
Nopoint
28 Ø 3.21
Condition
By Blocktype

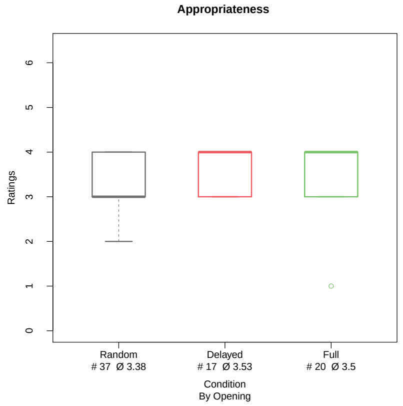

Appropriateness
Overall
Count
0
10
20
30
40
50
0
1
2
3
4
Rating

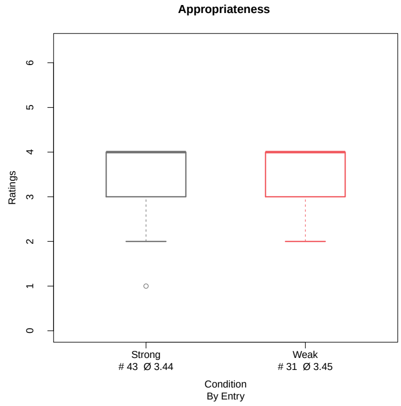

Appropriateness
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

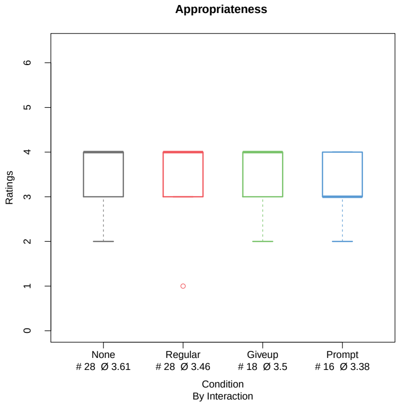

Appropriateness
Strong
Weak
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Entry

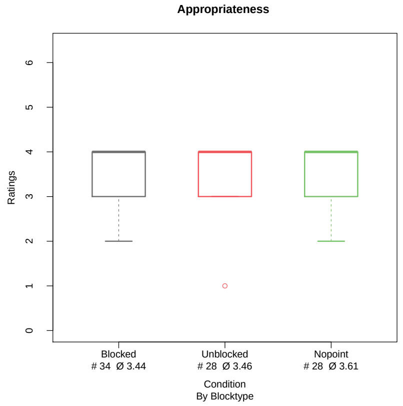

Appropriateness
None
Regular
Giveup
Prompt
Count
0
5
10
15
0
1
2
3
4
Rating
By Interaction

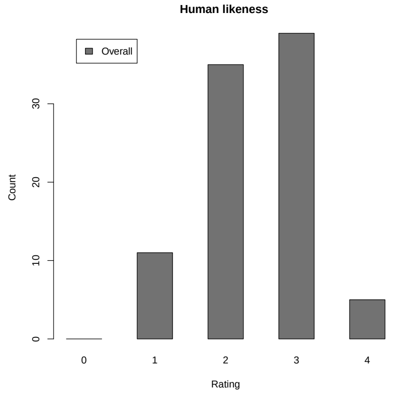

Appropriateness
Blocked
Unblocked
Nopoint
Count
0
5
10
15
0
1
2
3
4
Rating
By Blocktype

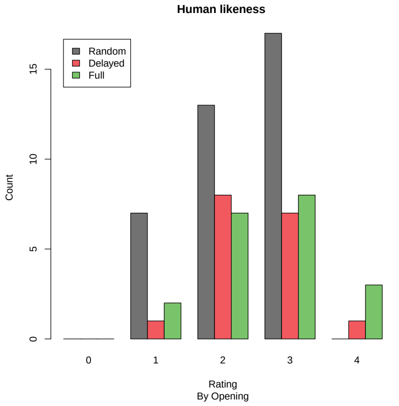

Appropriateness
Ratings
0
1
2
3
4
5
6
Random
37 Ø 3.38
Delayed
17 Ø 3.53
Full
20 Ø 3.5
Condition
By Opening

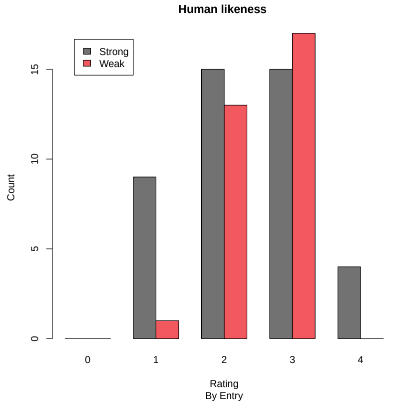

Appropriateness
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 3.44
Weak
31 Ø 3.45
Condition
By Entry

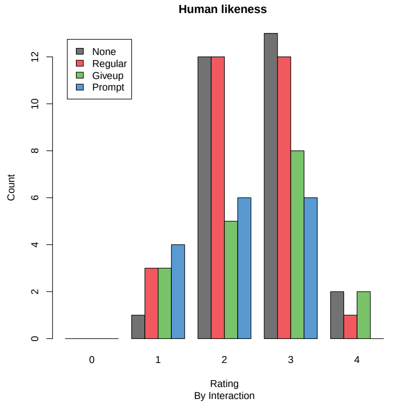

Appropriateness
Ratings
0
1
2
3
4
5
6
None
28 Ø 3.61
Regular
28 Ø 3.46
Giveup
18 Ø 3.5
Prompt
16 Ø 3.38
Condition
By Interaction

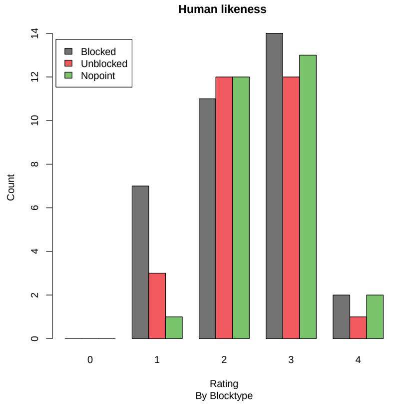

Appropriateness
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 3.44
Unblocked
28 Ø 3.46
Nopoint
28 Ø 3.61
Condition
By Blocktype

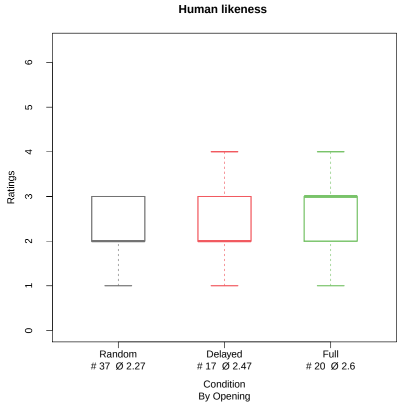

Human likeness
Overall
Count
0
10
20
30
0
1
2
3
4
Rating

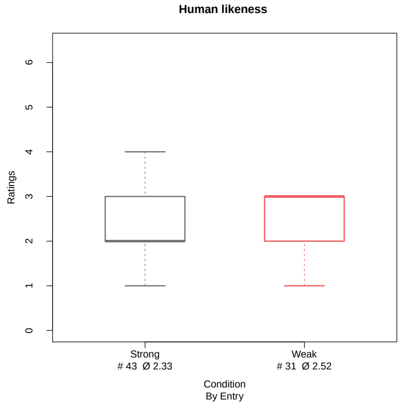

Human likeness
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

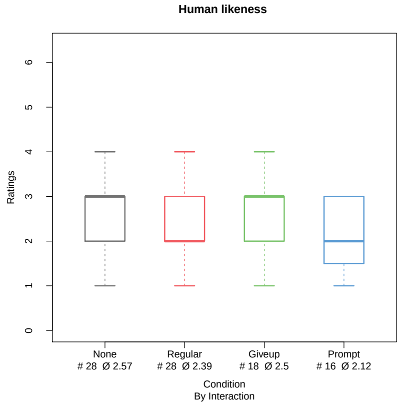

Human likeness
Strong
Weak
Count
0
5
10
15
0
1
2
3
4
Rating
By Entry

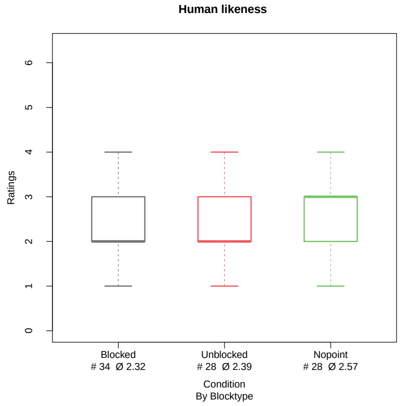

Human likeness
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Interaction

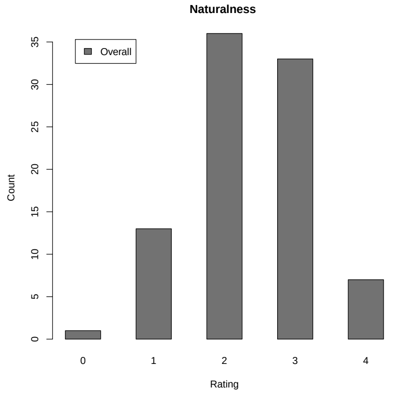

Human likeness
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Blocktype

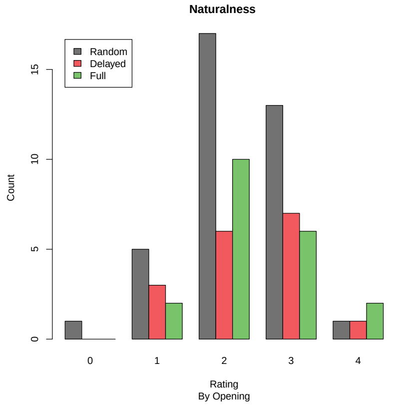

Human likeness
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.27
Delayed
17 Ø 2.47
Full
20 Ø 2.6
Condition
By Opening

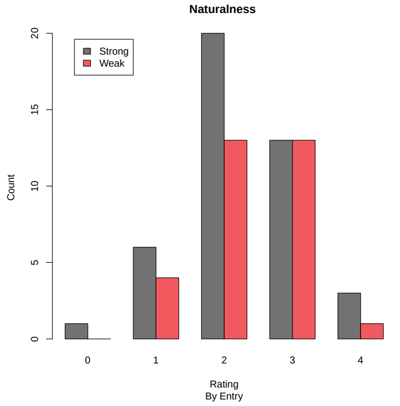

Human likeness
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.33
Weak
31 Ø 2.52
Condition
By Entry

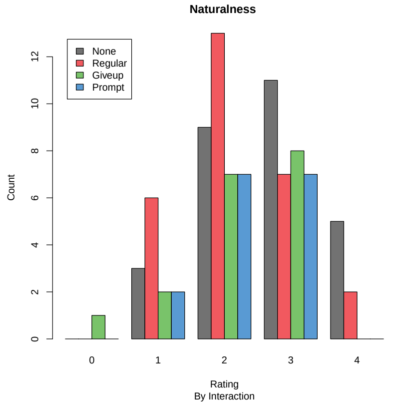

Human likeness
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.57
Regular
28 Ø 2.39
Giveup
18 Ø 2.5
Prompt
16 Ø 2.12
Condition
By Interaction

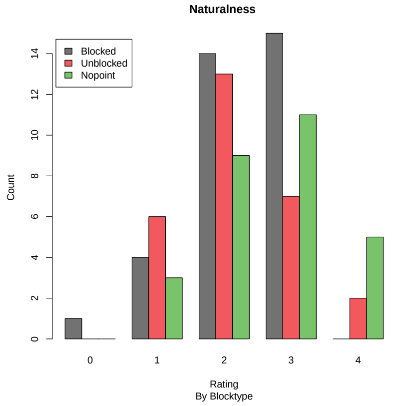

Human likeness
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.32
Unblocked
28 Ø 2.39
Nopoint
28 Ø 2.57
Condition
By Blocktype

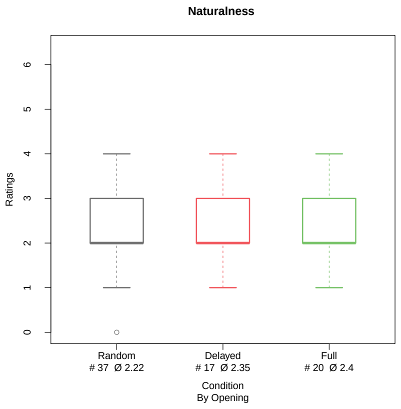

Naturalness
Overall
Count
0
5
10
15
20
25
30
35
0
1
2
3
4
Rating

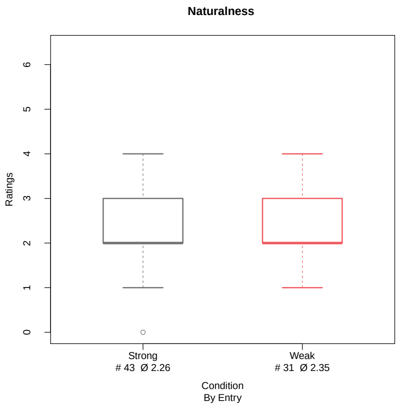

Naturalness
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

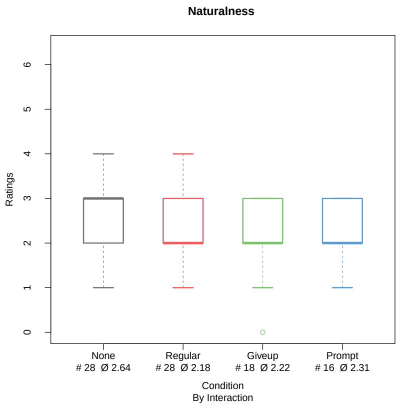

Naturalness
Strong
Weak
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Entry

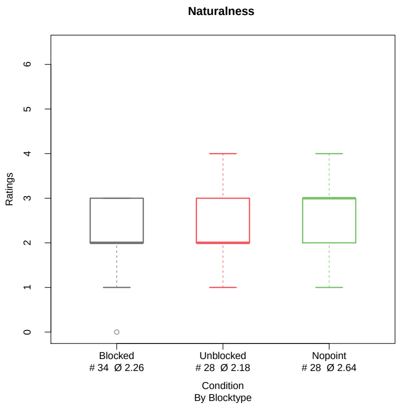

Naturalness
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Interaction

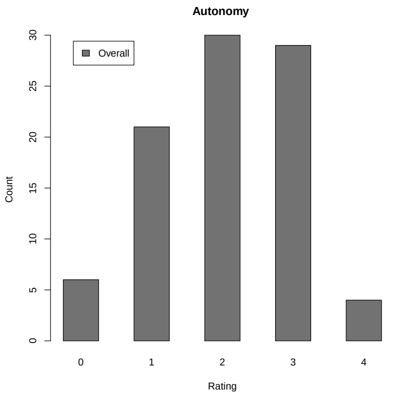

Naturalness
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Blocktype

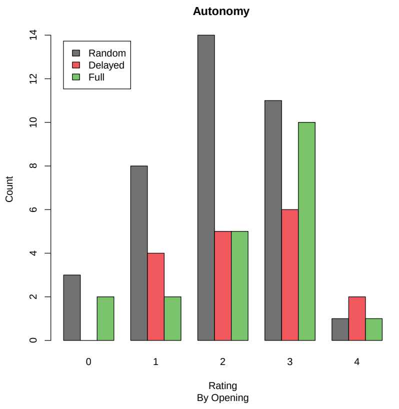

Naturalness
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.22
Delayed
17 Ø 2.35
Full
20 Ø 2.4
Condition
By Opening

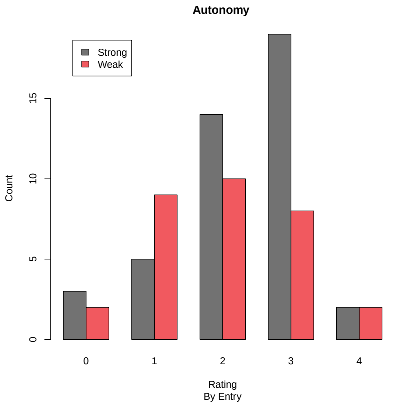

Naturalness
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.26
Weak
31 Ø 2.35
Condition
By Entry

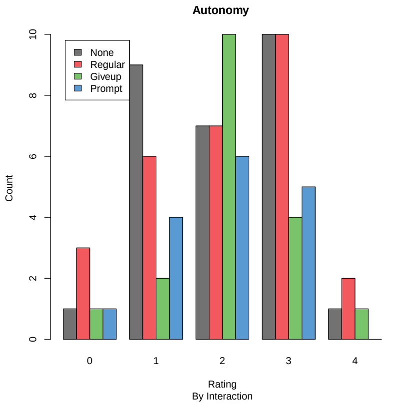

Naturalness
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.64
Regular
28 Ø 2.18
Giveup
18 Ø 2.22
Prompt
16 Ø 2.31
Condition
By Interaction

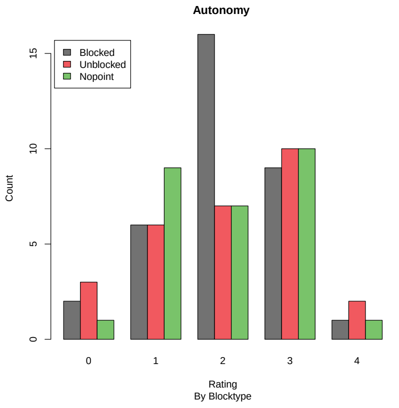

Naturalness
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.26
Unblocked
28 Ø 2.18
Nopoint
28 Ø 2.64
Condition
By Blocktype

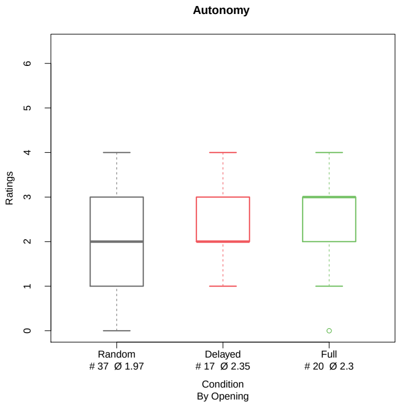

Autonomy
Overall
Count
0
5
10
15
20
25
30
0
1
2
3
4
Rating

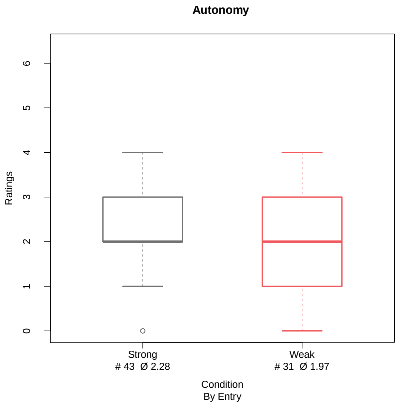

Autonomy
Random
Delayed
Full
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Opening

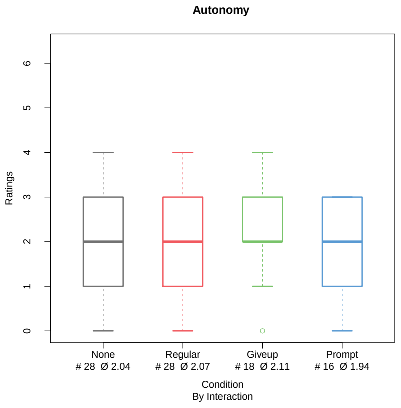

Autonomy
Strong
Weak
Count
0
5
10
15
0
1
2
3
4
Rating
By Entry

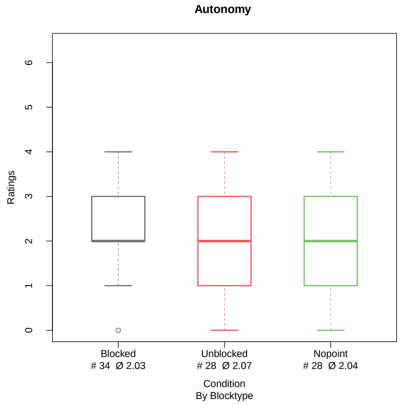

Autonomy
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
0
1
2
3
4
Rating
By Interaction

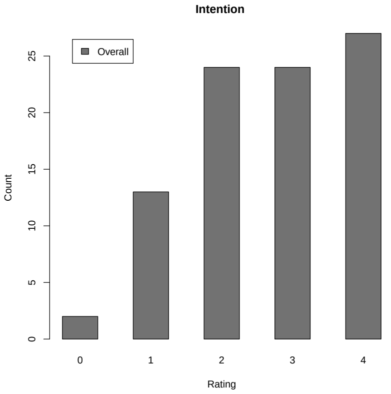

Autonomy
Blocked
Unblocked
Nopoint
Count
0
5
10
15
0
1
2
3
4
Rating
By Blocktype

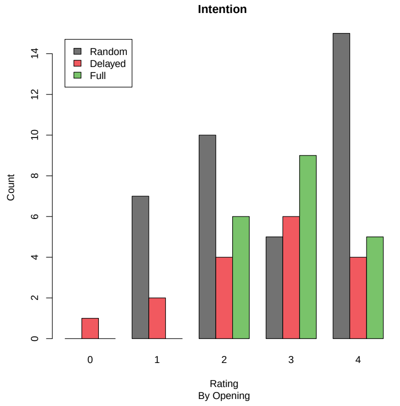

Autonomy
Ratings
0
1
2
3
4
5
6
Random
37 Ø 1.97
Delayed
17 Ø 2.35
Full
20 Ø 2.3
Condition
By Opening

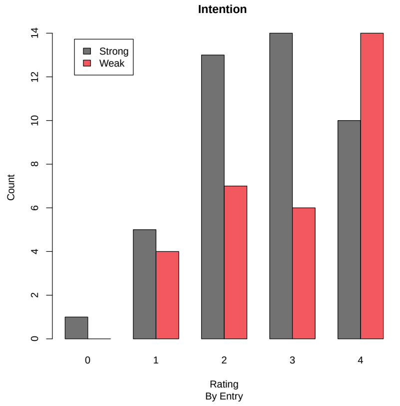

Autonomy
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.28
Weak
31 Ø 1.97
Condition
By Entry

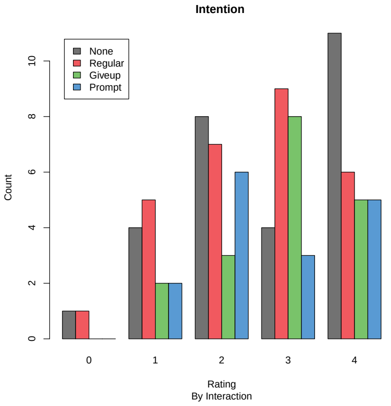

Autonomy
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.04
Regular
28 Ø 2.07
Giveup
18 Ø 2.11
Prompt
16 Ø 1.94
Condition
By Interaction

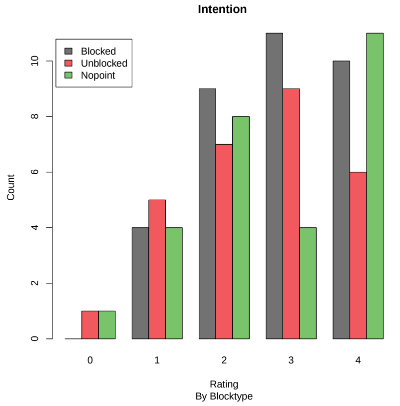

Autonomy
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.03
Unblocked
28 Ø 2.07
Nopoint
28 Ø 2.04
Condition
By Blocktype

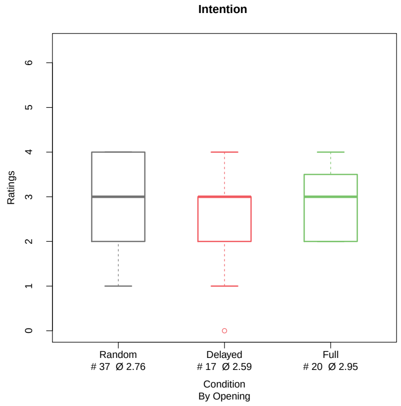

Intention
Overall
Count
0
5
10
15
20
25
0
1
2
3
4
Rating

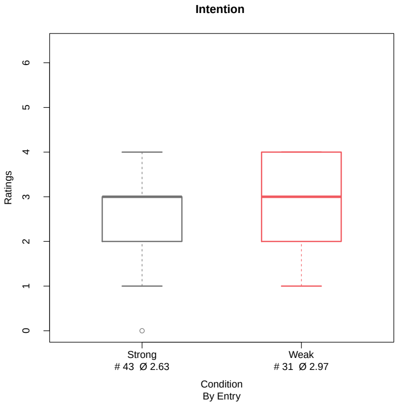

Intention
Random
Delayed
Full
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Opening

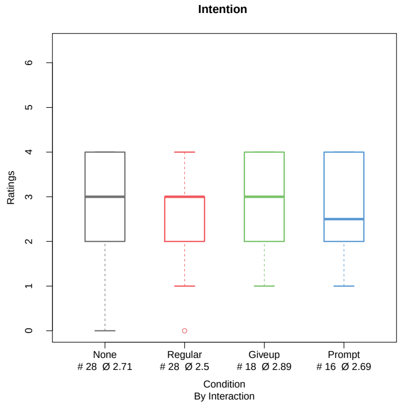

Intention
Strong
Weak
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Entry

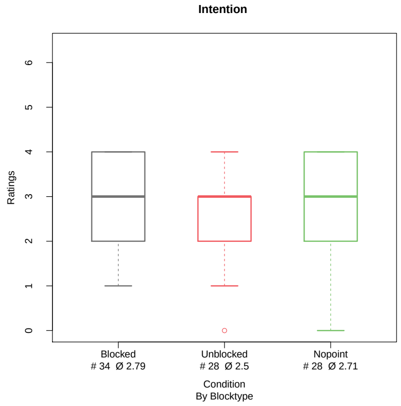

Intention
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
0
1
2
3
4
Rating
By Interaction

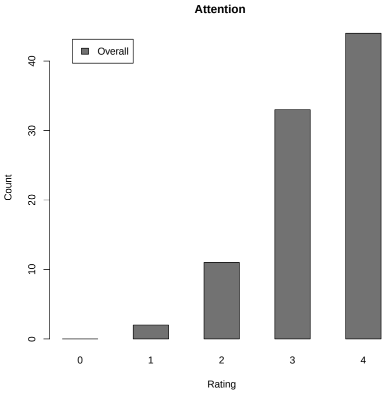

Intention
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
0
1
2
3
4
Rating
By Blocktype

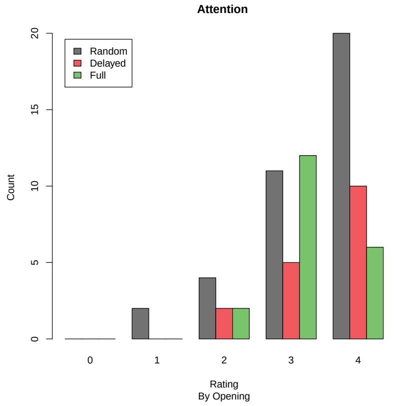

Intention
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.76
Delayed
17 Ø 2.59
Full
20 Ø 2.95
Condition
By Opening

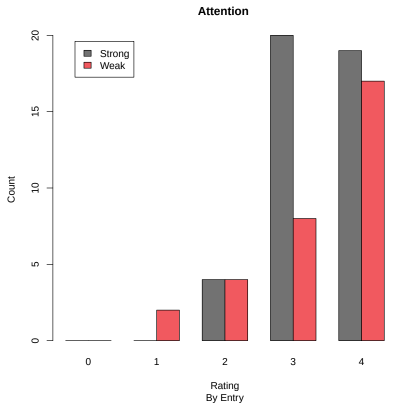

Intention
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.63
Weak
31 Ø 2.97
Condition
By Entry

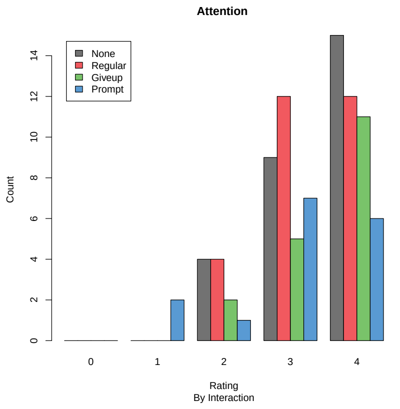

Intention
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.71
Regular
28 Ø 2.5
Giveup
18 Ø 2.89
Prompt
16 Ø 2.69
Condition
By Interaction

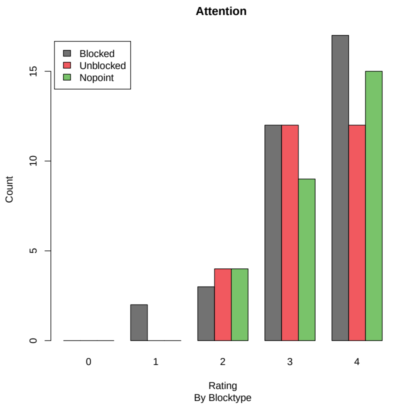

Intention
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.79
Unblocked
28 Ø 2.5
Nopoint
28 Ø 2.71
Condition
By Blocktype

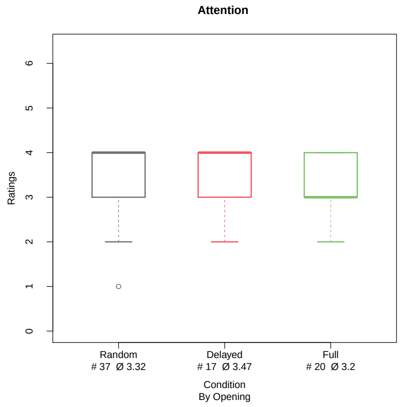

Attention
Overall
Count
0
10
20
30
40
0
1
2
3
4
Rating

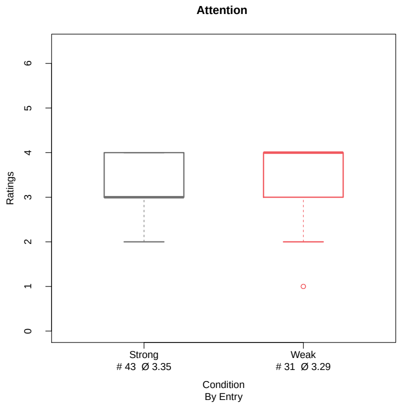

Attention
Random
Delayed
Full
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Opening

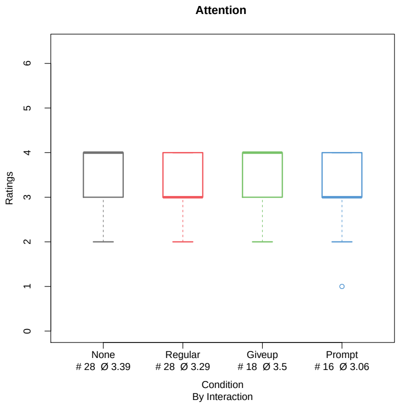

Attention
Strong
Weak
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Entry

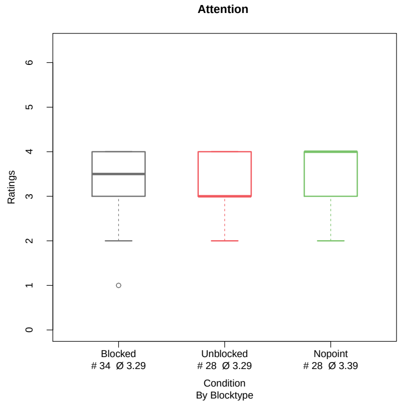

Attention
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Interaction

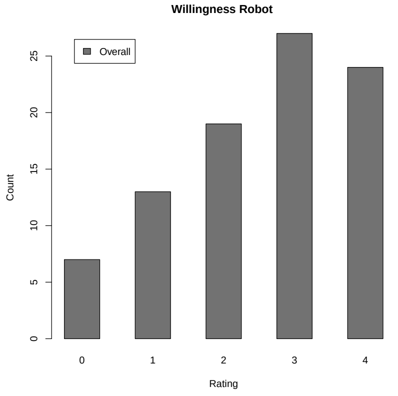

Attention
Blocked
Unblocked
Nopoint
Count
0
5
10
15
0
1
2
3
4
Rating
By Blocktype

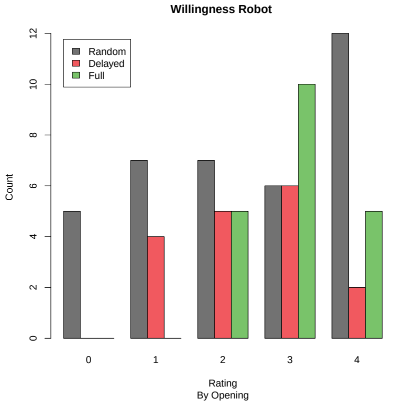

Attention
Ratings
0
1
2
3
4
5
6
Random
37 Ø 3.32
Delayed
17 Ø 3.47
Full
20 Ø 3.2
Condition
By Opening

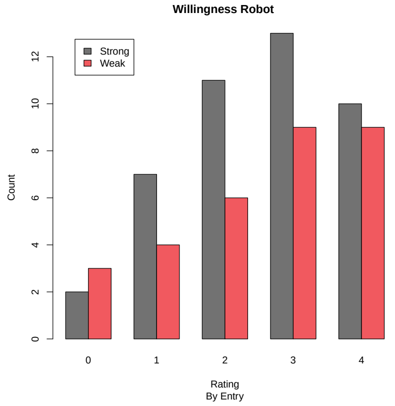

Attention
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 3.35
Weak
31 Ø 3.29
Condition
By Entry

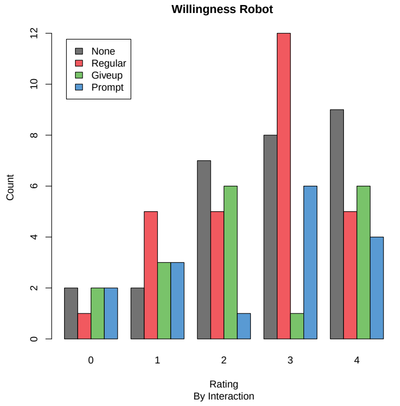

Attention
Ratings
0
1
2
3
4
5
6
None
28 Ø 3.39
Regular
28 Ø 3.29
Giveup
18 Ø 3.5
Prompt
16 Ø 3.06
Condition
By Interaction

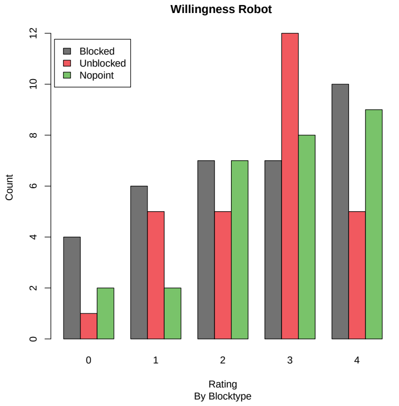

Attention
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 3.29
Unblocked
28 Ø 3.29
Nopoint
28 Ø 3.39
Condition
By Blocktype

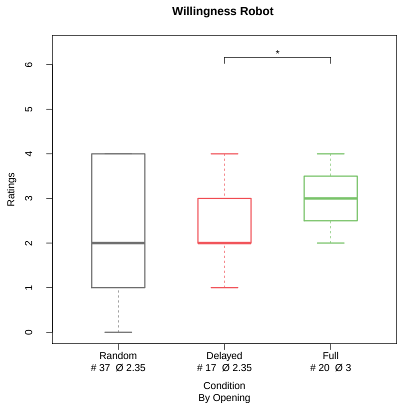

Willingness Robot
Overall
Count
0
5
10
15
20
25
0
1
2
3
4
Rating

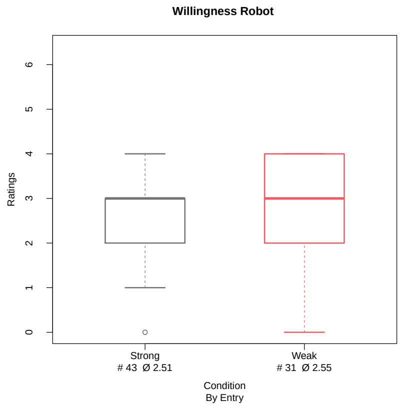

Willingness Robot
Random
Delayed
Full
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Opening

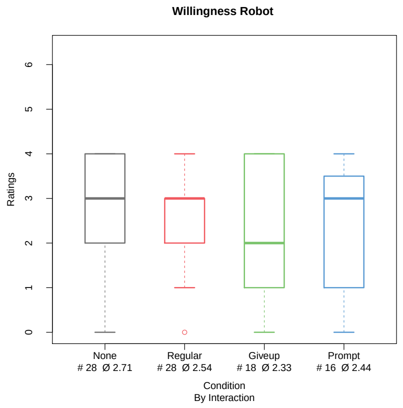

Willingness Robot
Strong
Weak
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Entry

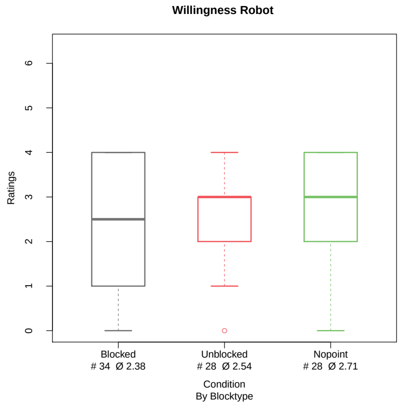

Willingness Robot
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Interaction

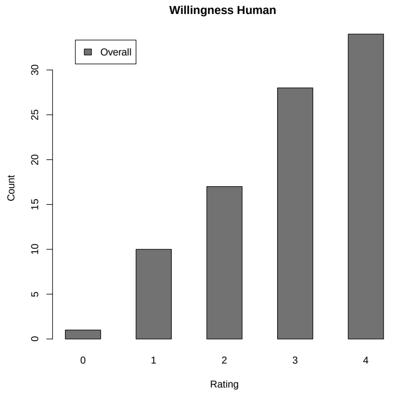

Willingness Robot
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Blocktype

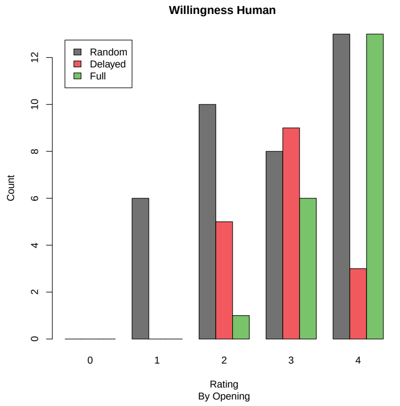

Willingness Robot
*
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.35
Delayed
17 Ø 2.35
Full
20 Ø 3
Condition
By Opening

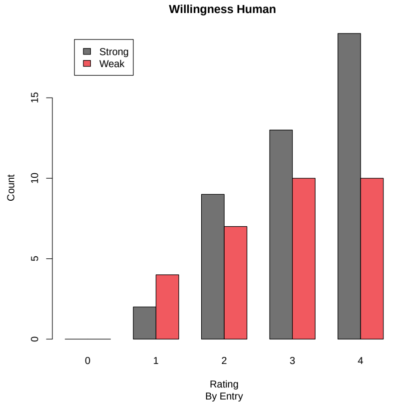

Willingness Robot
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.51
Weak
31 Ø 2.55
Condition
By Entry

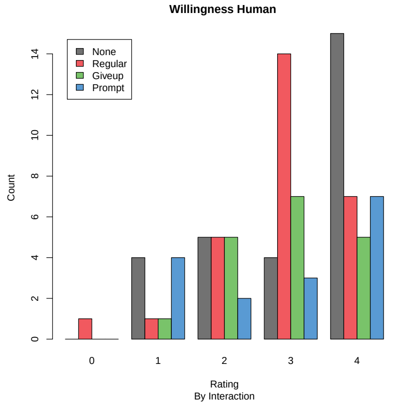

Willingness Robot
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.71
Regular
28 Ø 2.54
Giveup
18 Ø 2.33
Prompt
16 Ø 2.44
Condition
By Interaction

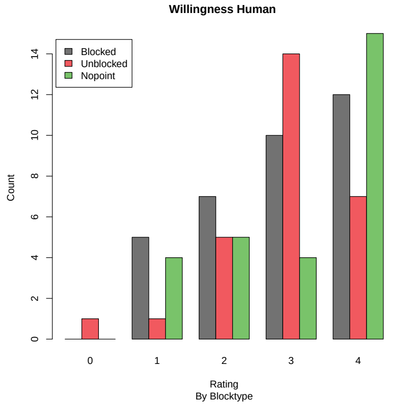

Willingness Robot
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.38
Unblocked
28 Ø 2.54
Nopoint
28 Ø 2.71
Condition
By Blocktype

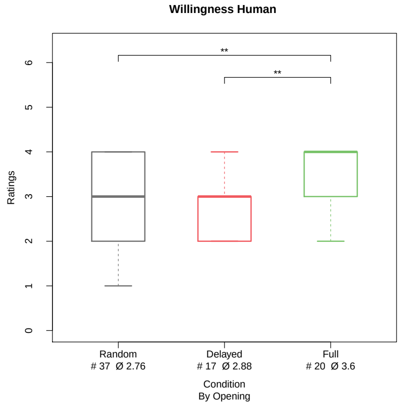

Willingness Human
Overall
Count
0
5
10
15
20
25
30
0
1
2
3
4
Rating

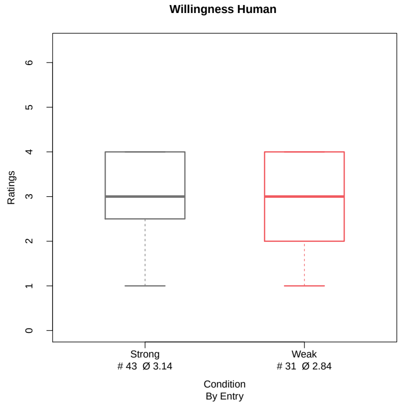

Willingness Human
Random
Delayed
Full
Count
0
2
4
6
8
10
12
0
1
2
3
4
Rating
By Opening

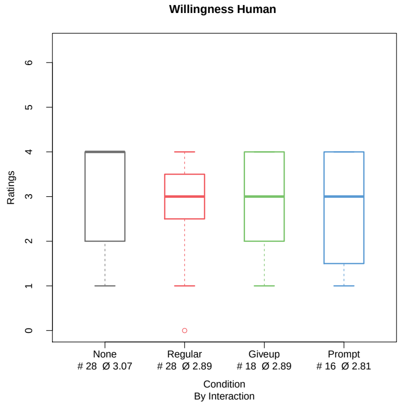

Willingness Human
Strong
Weak
Count
0
5
10
15
0
1
2
3
4
Rating
By Entry

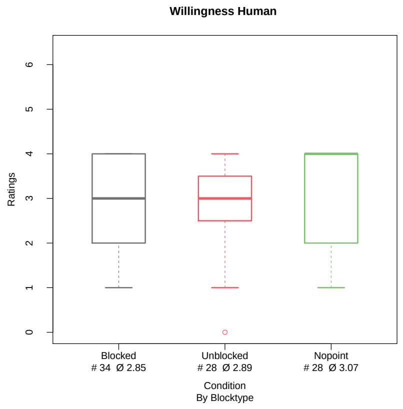

Willingness Human
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Interaction

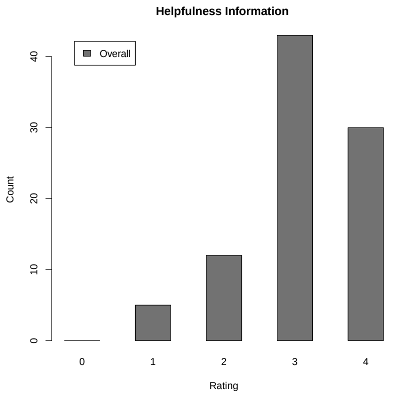

Willingness Human
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Blocktype

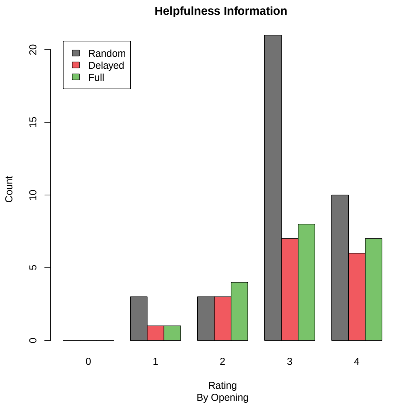

Willingness Human
**
**
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.76
Delayed
17 Ø 2.88
Full
20 Ø 3.6
Condition
By Opening

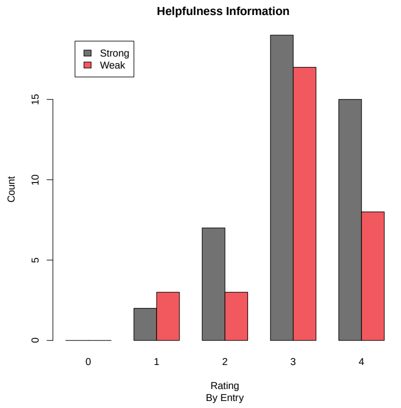

Willingness Human
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 3.14
Weak
31 Ø 2.84
Condition
By Entry

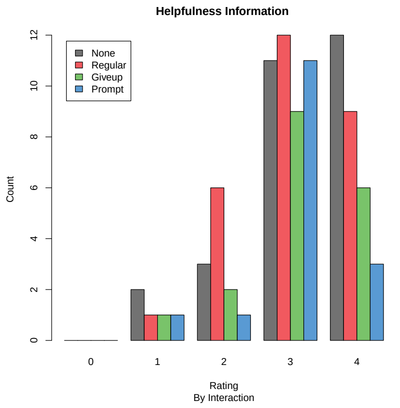

Willingness Human
Ratings
0
1
2
3
4
5
6
None
28 Ø 3.07
Regular
28 Ø 2.89
Giveup
18 Ø 2.89
Prompt
16 Ø 2.81
Condition
By Interaction

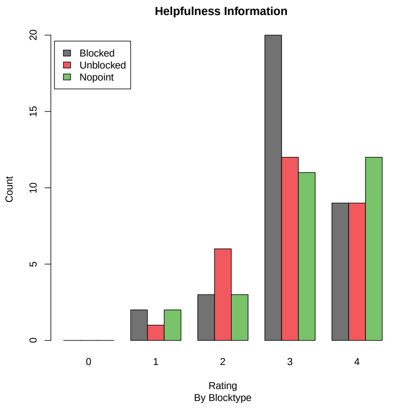

Willingness Human

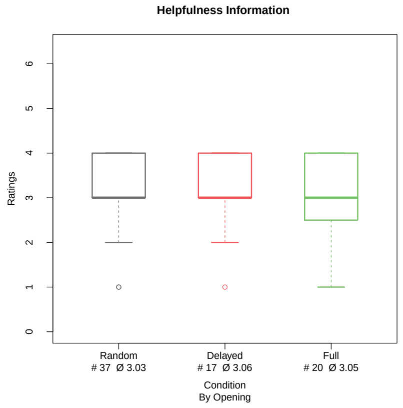

Helpfulness Information

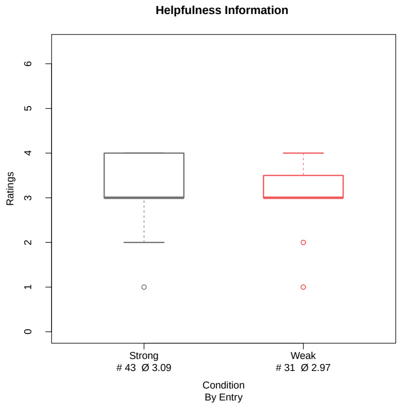

Helpfulness Information

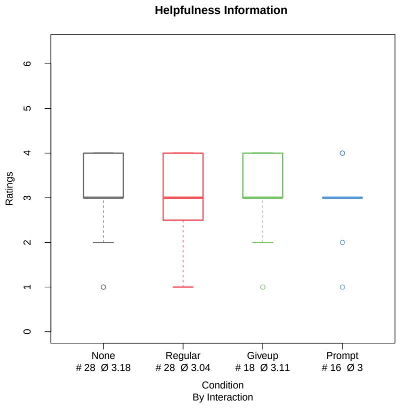

Helpfulness Information

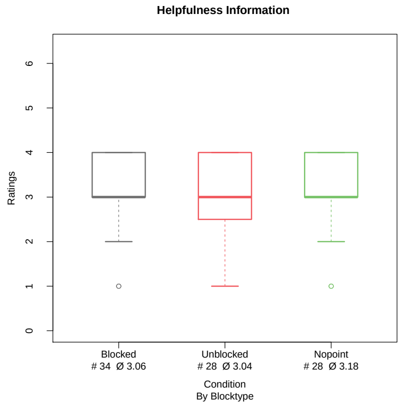

Helpfulness Information

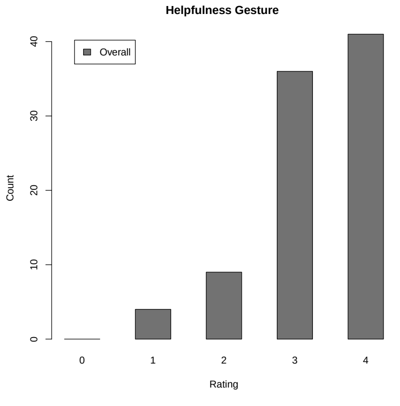

Helpfulness Information

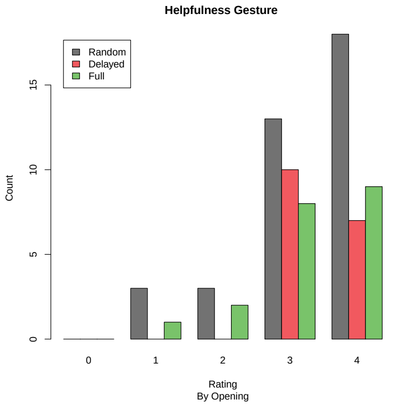

Helpfulness Information


Helpfulness Information

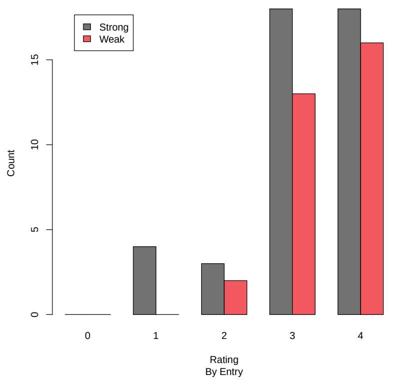

Helpfulness Information


Helpfulness Information

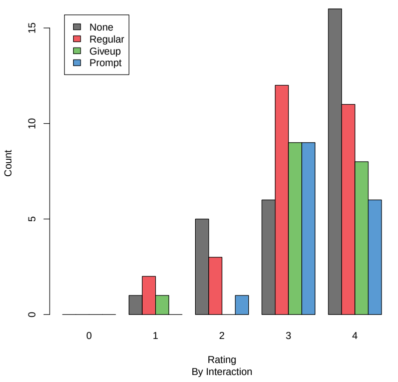

Helpfulness Gesture


Helpfulness Gesture

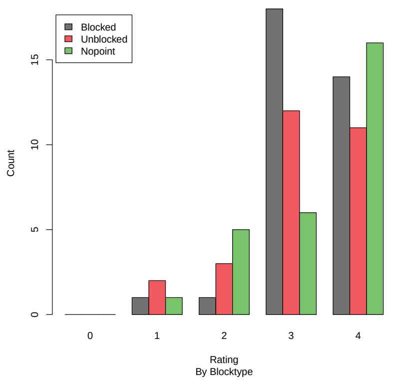
Helpfulness Gesture



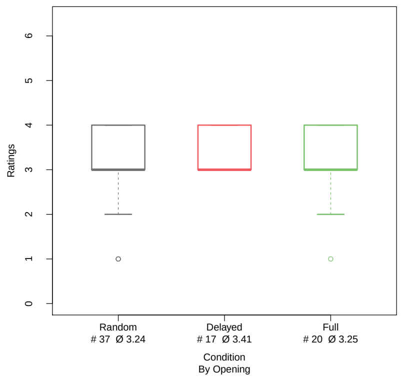
Helpfulness Gesture



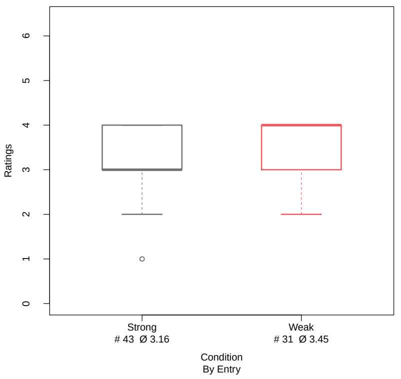
Helpfulness Gesture



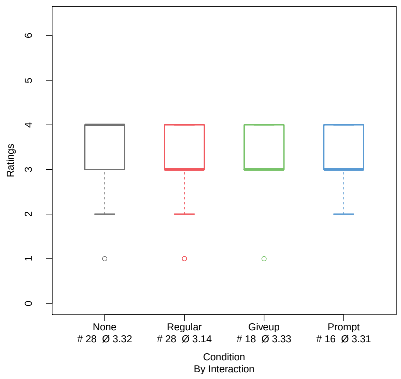
Helpfulness Gesture



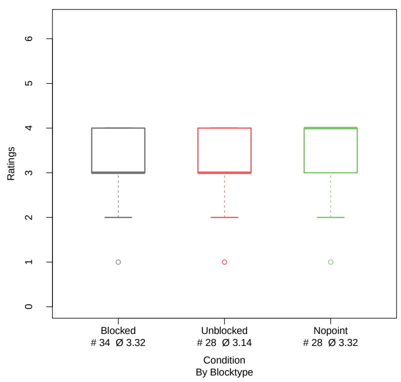
Helpfulness Gesture



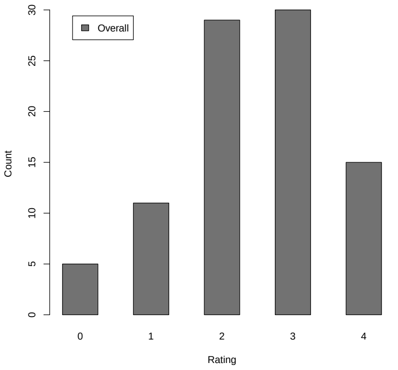
Helpfulness Gesture



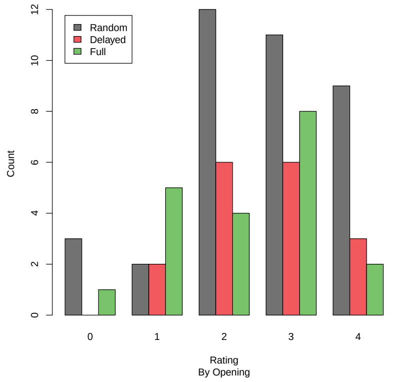
Helpfulness Gesture



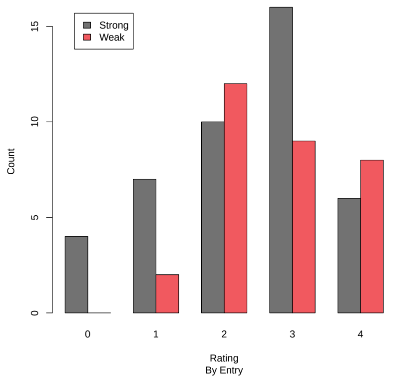
Helpfulness Gaze



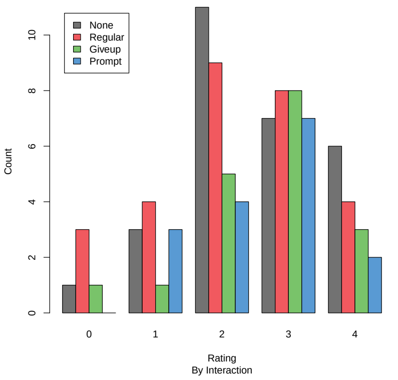
Helpfulness Gaze



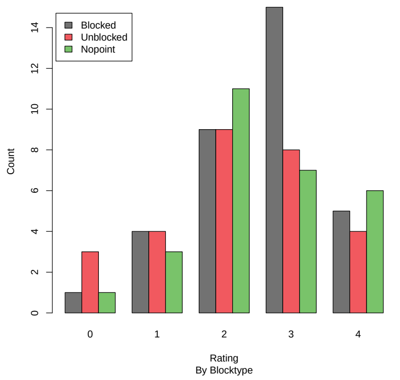
Helpfulness Gaze

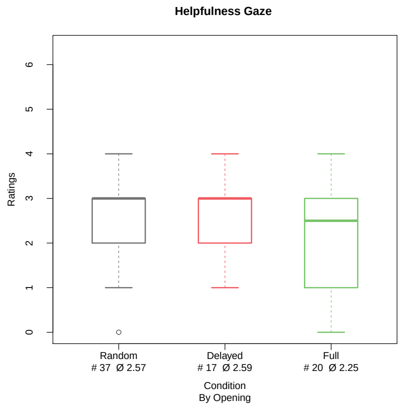


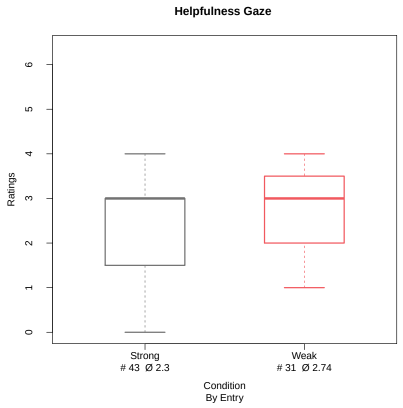
Helpfulness Gaze

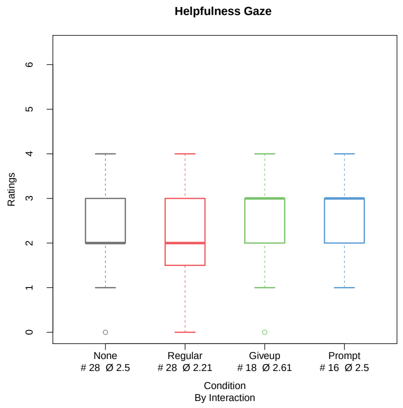


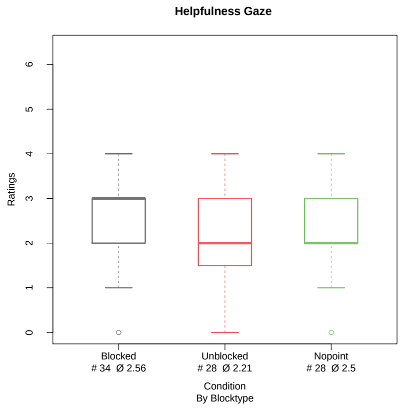
Helpfulness Gaze

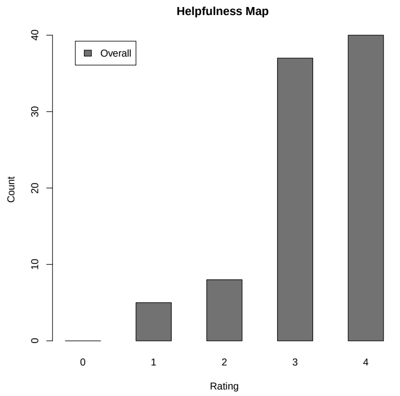


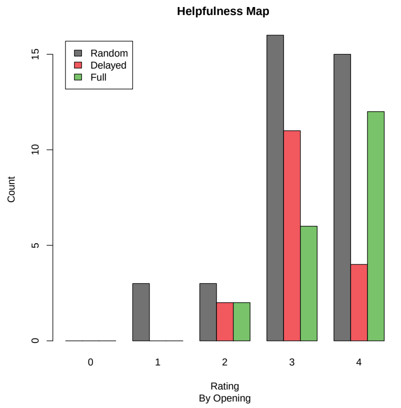

Helpfulness Gaze
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.57
Delayed
17 Ø 2.59
Full
20 Ø 2.25
Condition
By Opening

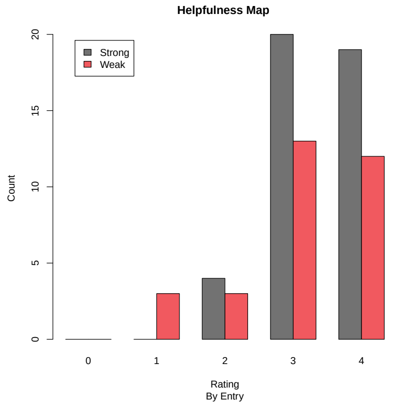

Helpfulness Gaze
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.3
Weak
31 Ø 2.74
Condition
By Entry

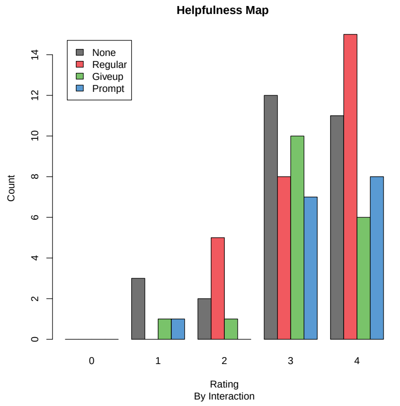

Helpfulness Gaze
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.5
Regular
28 Ø 2.21
Giveup
18 Ø 2.61
Prompt
16 Ø 2.5
Condition
By Interaction

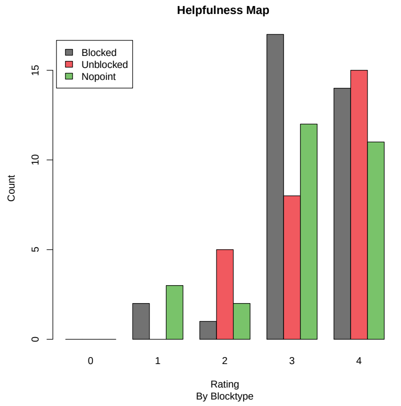

Helpfulness Gaze
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.56
Unblocked
28 Ø 2.21
Nopoint
28 Ø 2.5
Condition
By Blocktype

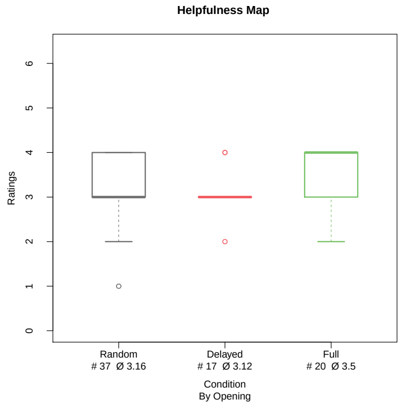

Helpfulness Map
Overall
Count
0
10
20
30
40
0
1
2
3
4
Rating

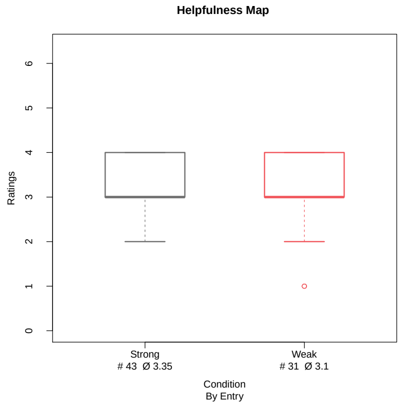

Helpfulness Map
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

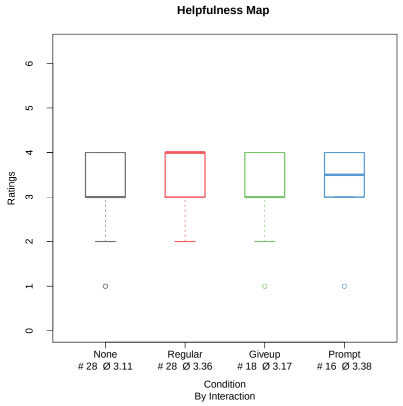

Helpfulness Map
Strong
Weak
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Entry

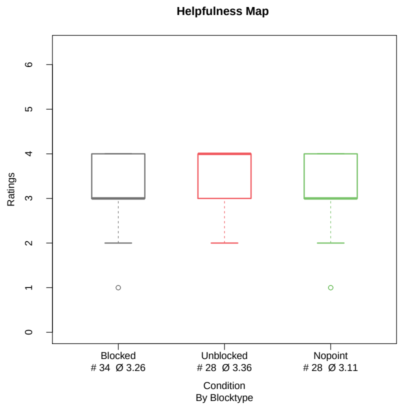

Helpfulness Map
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Interaction

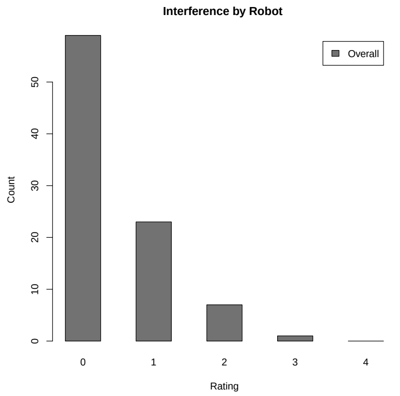

Helpfulness Map
Blocked
Unblocked
Nopoint
Count
0
5
10
15
0
1
2
3
4
Rating
By Blocktype

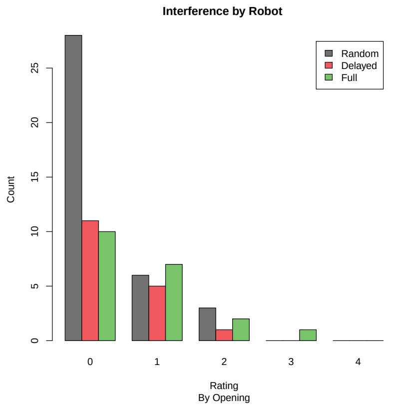

Helpfulness Map
Ratings
0
1
2
3
4
5
6
Random
37 Ø 3.16
Delayed
17 Ø 3.12
Full
20 Ø 3.5
Condition
By Opening

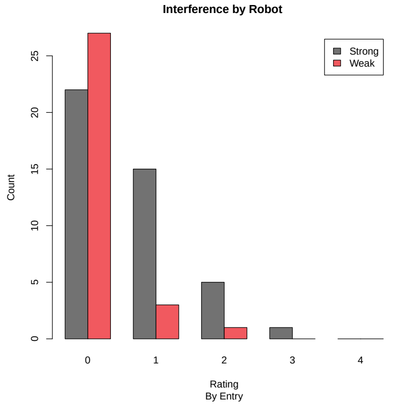

Helpfulness Map
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 3.35
Weak
31 Ø 3.1
Condition
By Entry

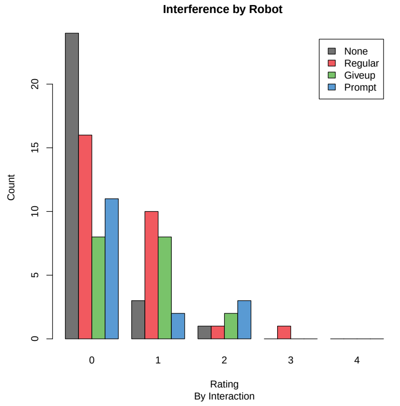

Helpfulness Map
Ratings
0
1
2
3
4
5
6
None
28 Ø 3.11
Regular
28 Ø 3.36
Giveup
18 Ø 3.17
Prompt
16 Ø 3.38
Condition
By Interaction

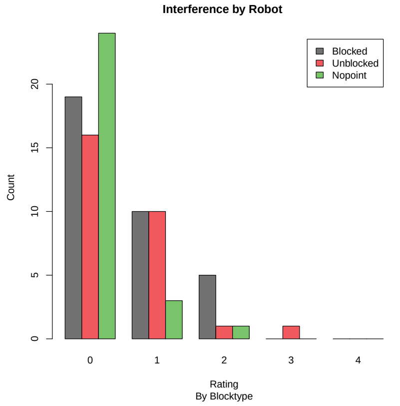

Helpfulness Map
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 3.26
Unblocked
28 Ø 3.36
Nopoint
28 Ø 3.11
Condition
By Blocktype

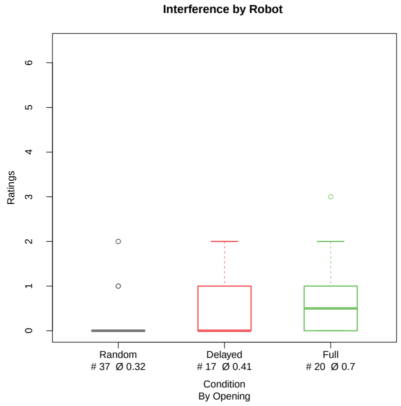

Interference by Robot
Overall
Count
0
10
20
30
40
50
0
1
2
3
4
Rating

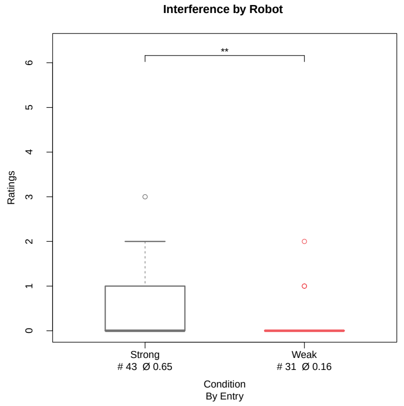

Interference by Robot
Random
Delayed
Full
Count
0
5
10
15
20
25
0
1
2
3
4
Rating
By Opening

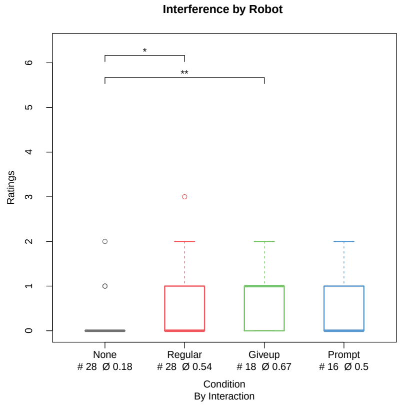

Interference by Robot
Strong
Weak
Count
0
5
10
15
20
25
0
1
2
3
4
Rating
By Entry

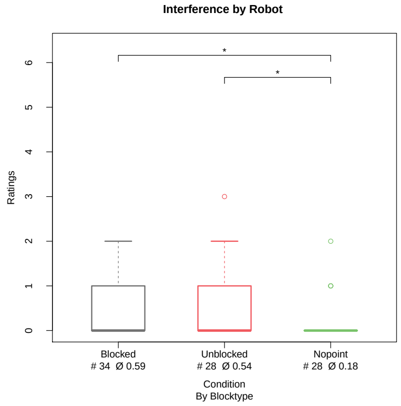

Interference by Robot
None
Regular
Giveup
Prompt
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Interaction

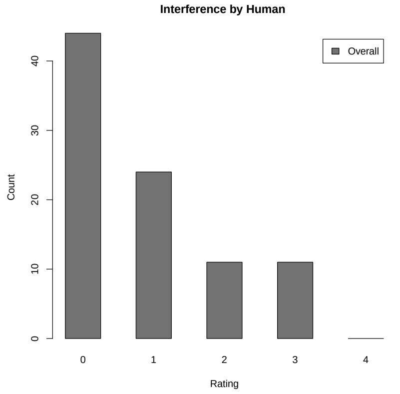

Interference by Robot
Blocked
Unblocked
Nopoint
Count
0
5
10
15
20
0
1
2
3
4
Rating
By Blocktype

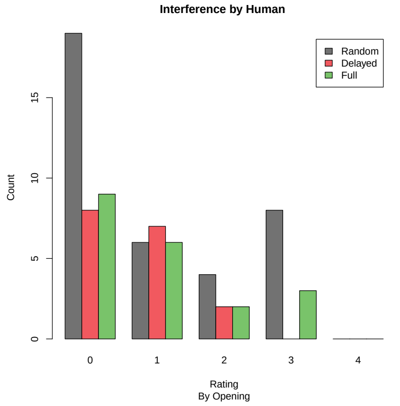

Interference by Robot
Ratings
0
1
2
3
4
5
6
Random
37 Ø 0.32
Delayed
17 Ø 0.41
Full
20 Ø 0.7
Condition
By Opening

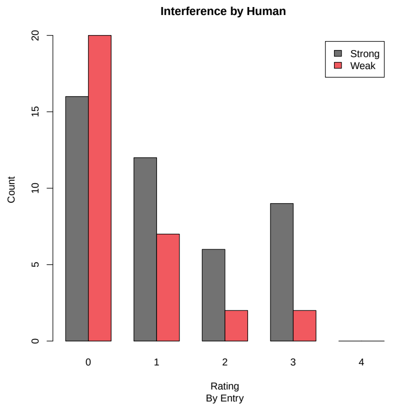

Interference by Robot
**
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 0.65
Weak
31 Ø 0.16
Condition
By Entry

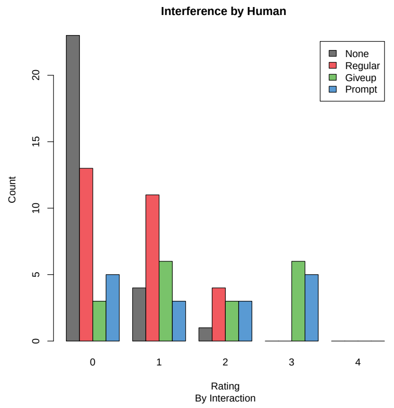

Interference by Robot
*
**
Ratings
0
1
2
3
4
5
6
None
28 Ø 0.18
Regular
28 Ø 0.54
Giveup
18 Ø 0.67
Prompt
16 Ø 0.5
Condition
By Interaction

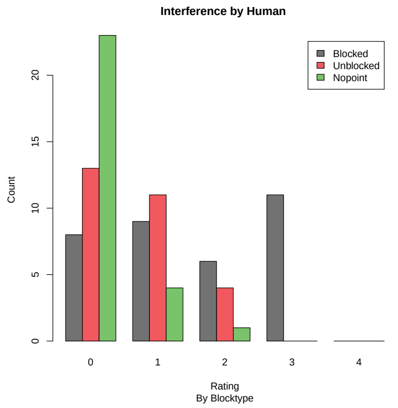

Interference by Robot
*
*
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 0.59
Unblocked
28 Ø 0.54
Nopoint
28 Ø 0.18
Condition
By Blocktype

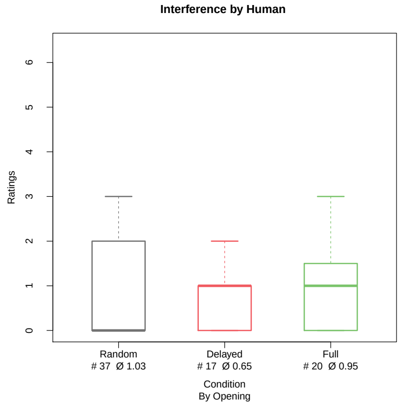

Interference by Human
Overall
Count
0
10
20
30
40
0
1
2
3
4
Rating

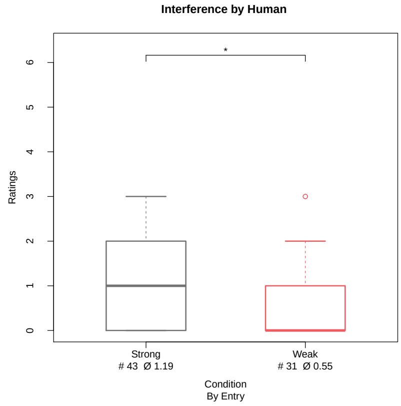

Interference by Human
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

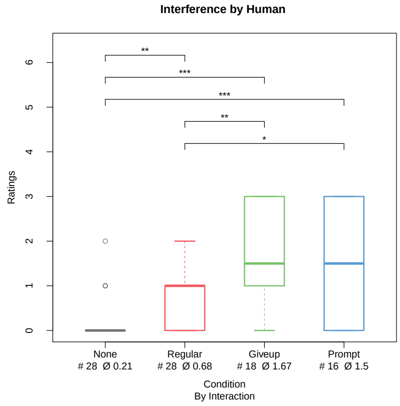

Interference by Human

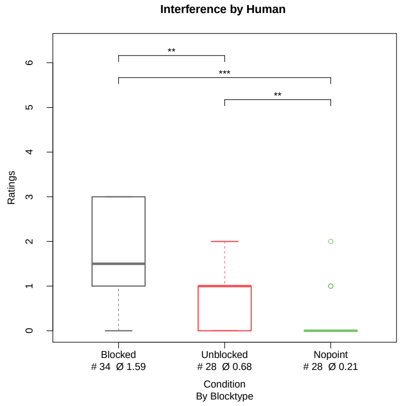

Interference by Human

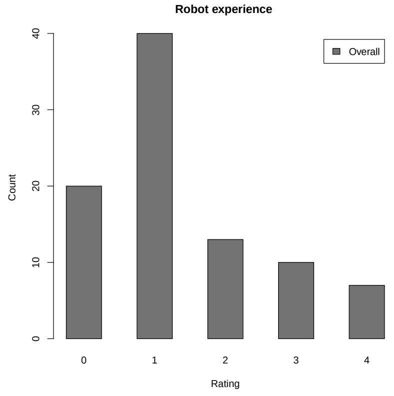

Interference by Human

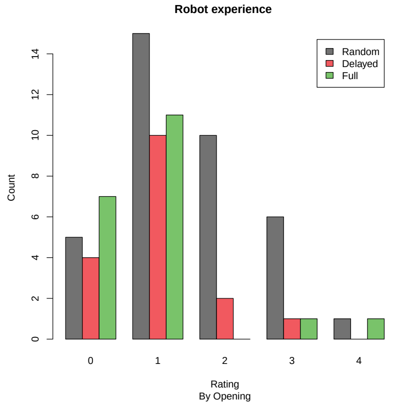

Interference by Human

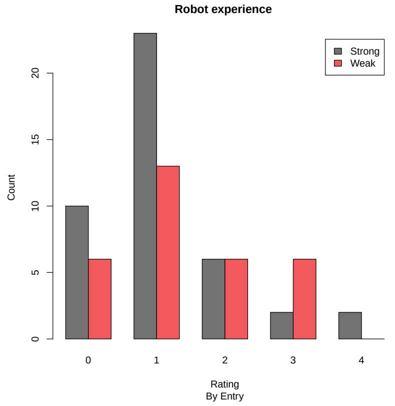

Interference by Human

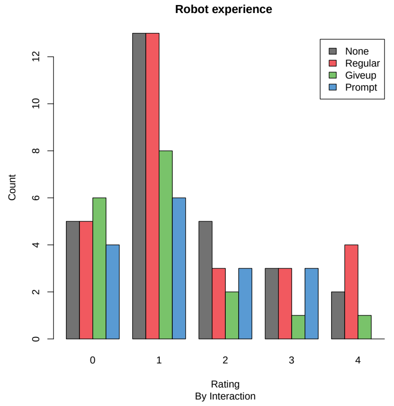

Interference by Human

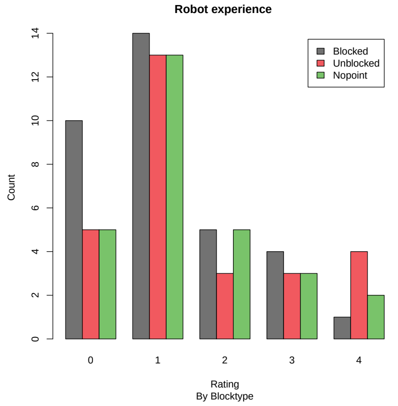

Interference by Human

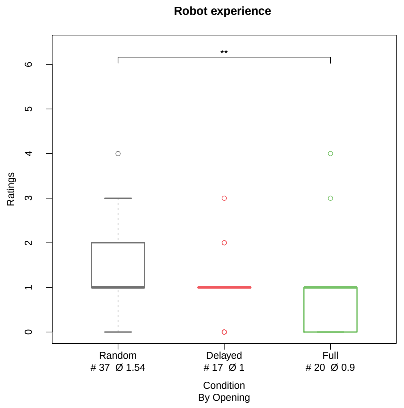

Robot experience

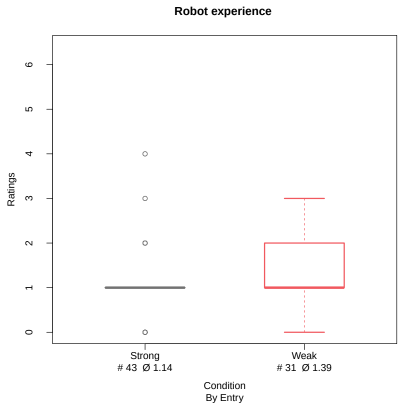

Robot experience

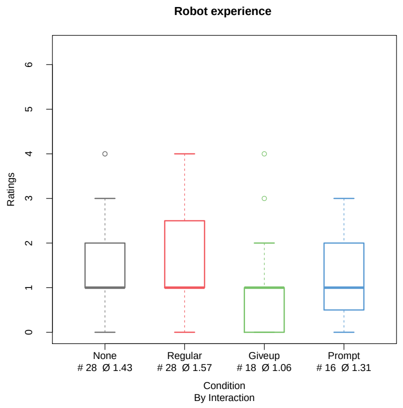

Robot experience

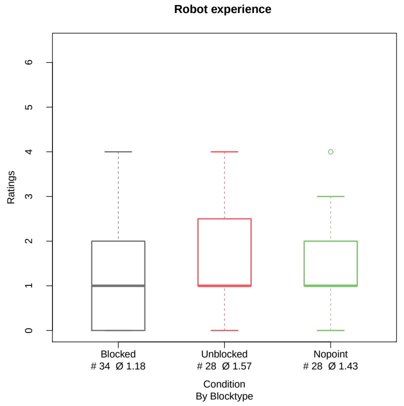

Robot experience

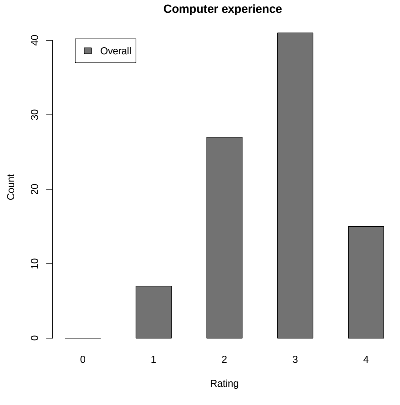

Robot experience

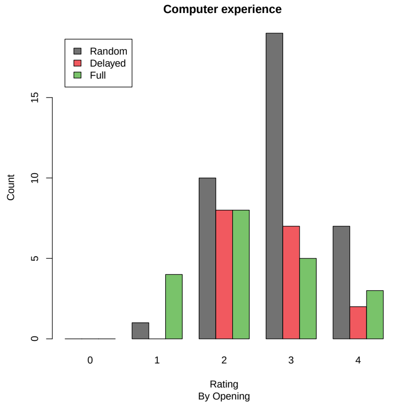

Robot experience
**
Ratings
0
1
2
3
4
5
6
Random
37 Ø 1.54
Delayed
17 Ø 1
Full
20 Ø 0.9
Condition
By Opening

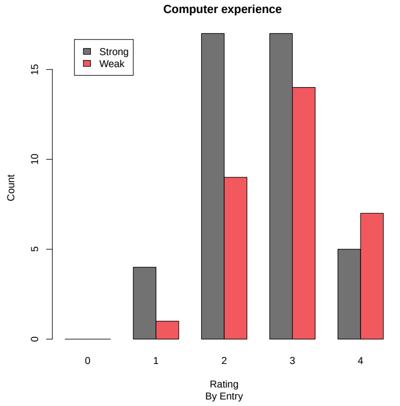

Robot experience
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 1.14
Weak
31 Ø 1.39
Condition
By Entry

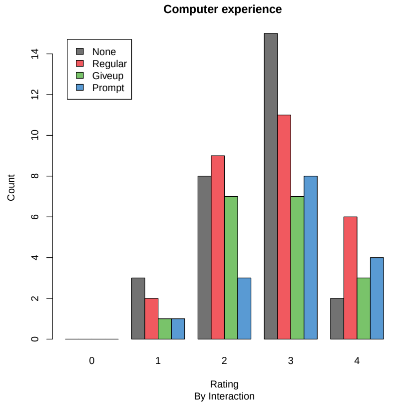

Robot experience
Ratings
0
1
2
3
4
5
6
None
28 Ø 1.43
Regular
28 Ø 1.57
Giveup
18 Ø 1.06
Prompt
16 Ø 1.31
Condition
By Interaction

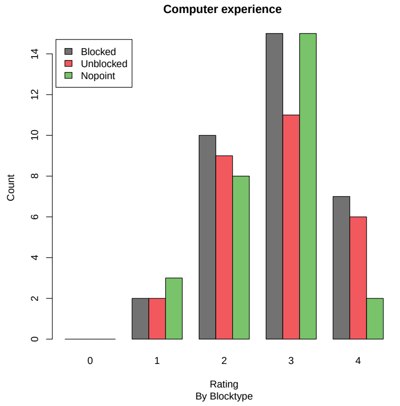

Robot experience
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 1.18
Unblocked
28 Ø 1.57
Nopoint
28 Ø 1.43
Condition
By Blocktype

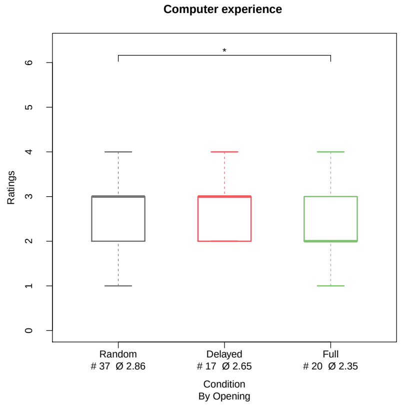

Computer experience
Overall
Count
0
10
20
30
40
0
1
2
3
4
Rating

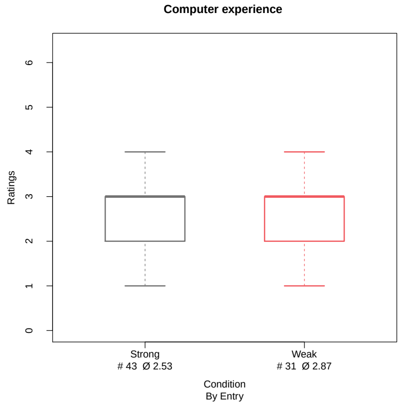

Computer experience
Random
Delayed
Full
Count
0
5
10
15
0
1
2
3
4
Rating
By Opening

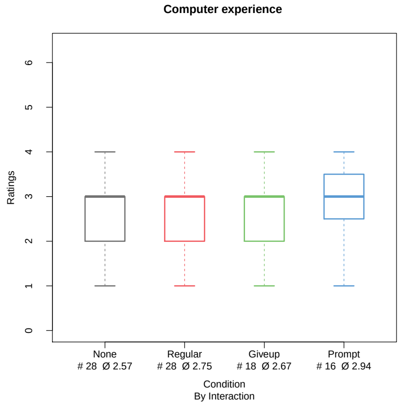

Computer experience
Strong
Weak
Count
0
5
10
15
0
1
2
3
4
Rating
By Entry

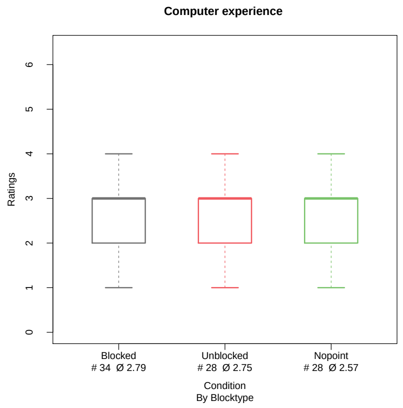

Computer experience
None
Regular
Giveup
Prompt
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Interaction

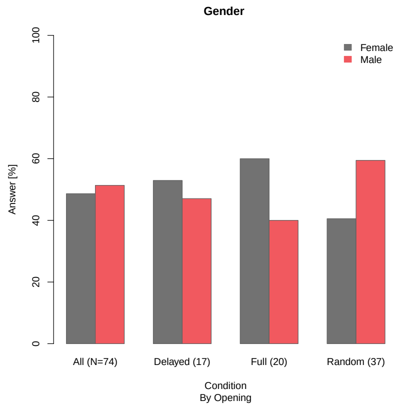

Computer experience
Blocked
Unblocked
Nopoint
Count
0
2
4
6
8
10
12
14
0
1
2
3
4
Rating
By Blocktype

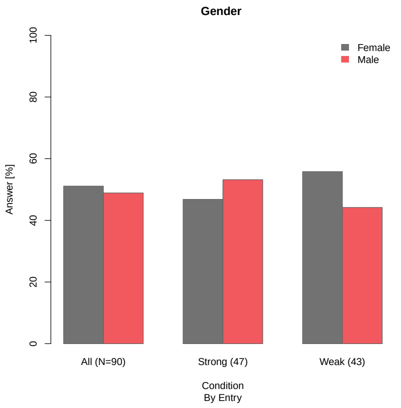

Computer experience
*
Ratings
0
1
2
3
4
5
6
Random
37 Ø 2.86
Delayed
17 Ø 2.65
Full
20 Ø 2.35
Condition
By Opening

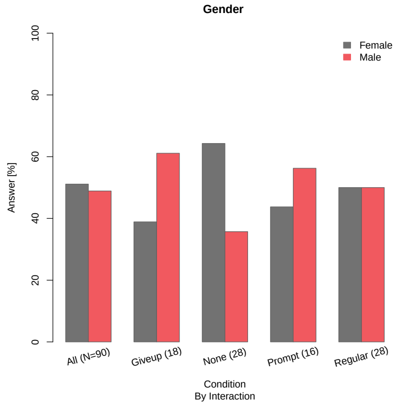

Computer experience
Ratings
0
1
2
3
4
5
6
Strong
43 Ø 2.53
Weak
31 Ø 2.87
Condition
By Entry

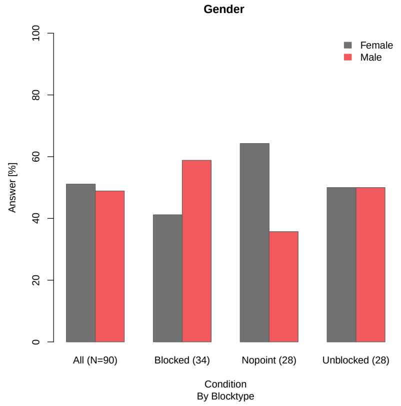

Computer experience
Ratings
0
1
2
3
4
5
6
None
28 Ø 2.57
Regular
28 Ø 2.75
Giveup
18 Ø 2.67
Prompt
16 Ø 2.94
Condition
By Interaction

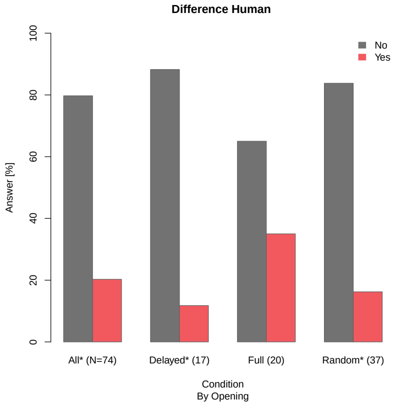

Computer experience
Ratings
0
1
2
3
4
5
6
Blocked
34 Ø 2.79
Unblocked
28 Ø 2.75
Nopoint
28 Ø 2.57
Condition
By Blocktype

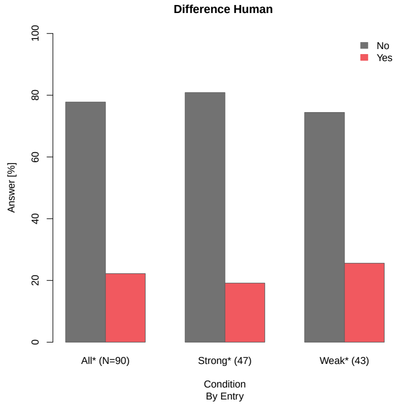

Gender
Answer [%]
0
20
40
60
80
100
Female
Male
All (N=74)
Delayed (17)
Full (20)
Random (37)
Condition
By Opening

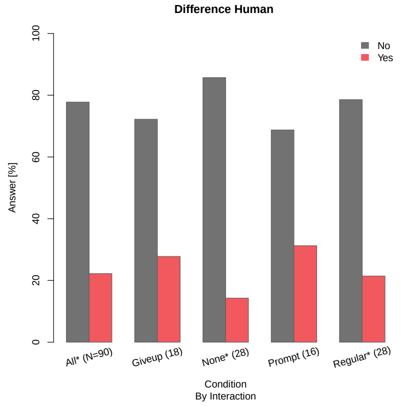

Gender
Answer [%]
0
20
40
60
80
100
Female
Male
All (N=90)
Strong (47)
Weak (43)
Condition
By Entry

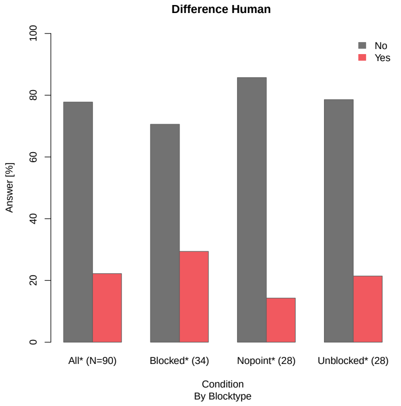

Gender
Answer [%]
0
20
40
60
80
100
Female
Male
All (N=90)
Giveup (18)
None (28)
Prompt (16)
Regular (28)
Condition
By Interaction

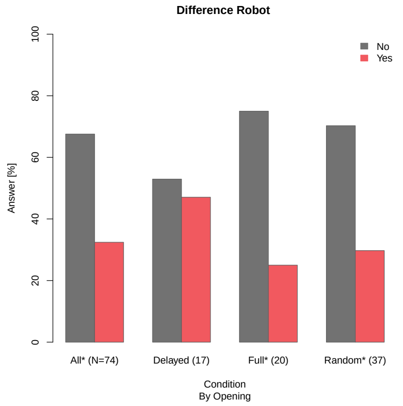

Gender
Answer [%]
0
20
40
60
80
100
Female
Male
All (N=90)
Blocked (34)
Nopoint (28)
Unblocked (28)
Condition
By Blocktype

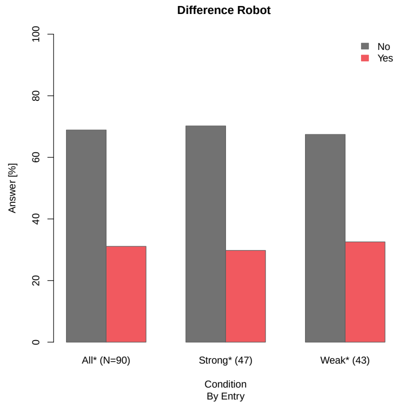

Difference Human
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=74)
Delayed* (17)
Full (20)
Random* (37)
Condition
By Opening

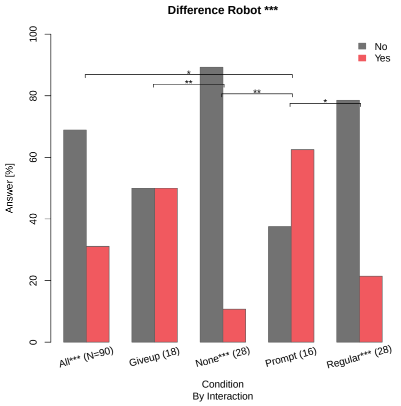

Difference Human
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=90)
Strong* (47)
Weak* (43)
Condition
By Entry

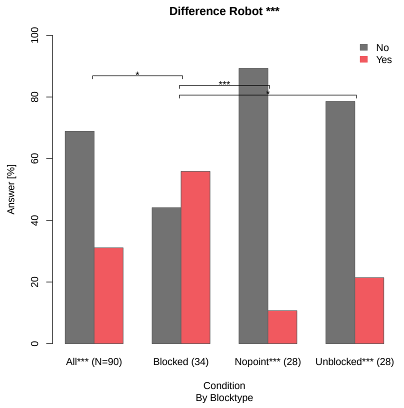

Difference Human
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=90)
Giveup (18)
None* (28)
Prompt (16)
Regular* (28)
Condition
By Interaction


Difference Human
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=90)
Blocked* (34)
Nopoint* (28)
Unblocked* (28)
Condition
By Blocktype

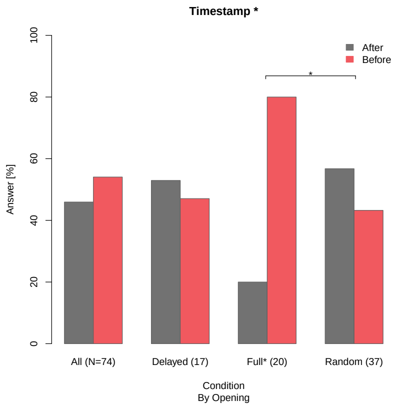

Difference Robot
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=74)
Delayed (17)
Full* (20)
Random* (37)
Condition
By Opening


Difference Robot
Answer [%]
0
20
40
60
80
100
No
Yes
All* (N=90)
Strong* (47)
Weak* (43)
Condition
By Entry

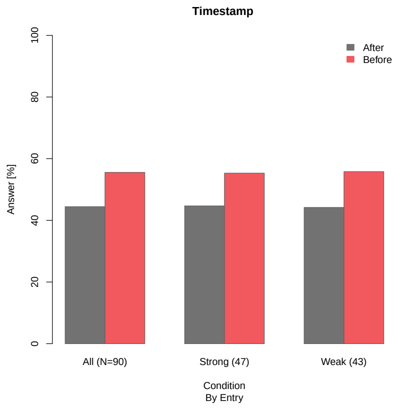

Difference Robot ***
Answer [%]
0
20
40
60
80
100
No
Yes
*
**
**
*
All*** (N=90)
Giveup (18)
None*** (28)
Prompt (16)
Regular*** (28)
Condition
By Interaction



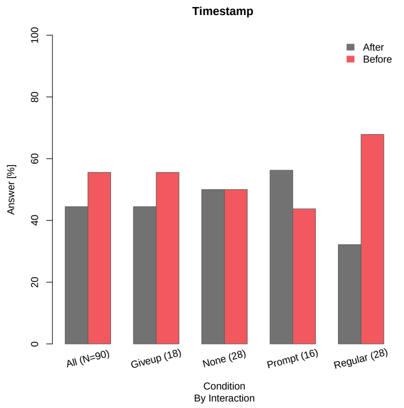
Difference Robot ***
Answer [%]
0
20
40
60
80
100
No
Yes
*
***
*
All*** (N=90)
Blocked (34)
Nopoint*** (28)
Unblocked*** (28)
Condition
By Blocktype



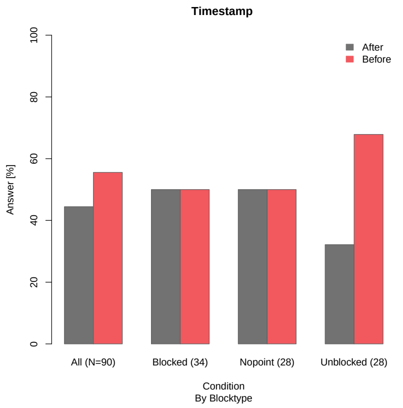
Timestamp *
Answer [%]
0
20
40
60
80
100
After
Before
*
All (N=74)
Delayed (17)
Full* (20)
Random (37)
Condition
By Opening



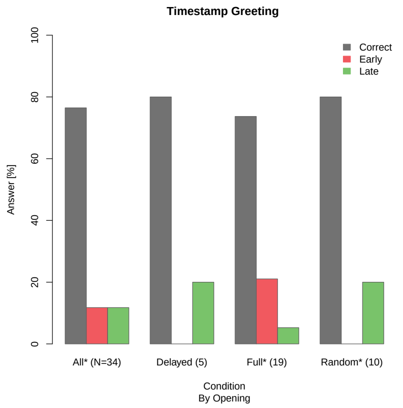
Timestamp
Answer [%]
0
20
40
60
80
100
After
Before
All (N=90)
Strong (47)
Weak (43)
Condition
By Entry



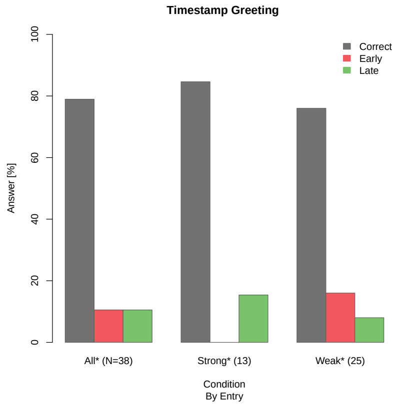
Timestamp
Answer [%]
0
20
40
60
80
100
After
Before
All (N=90)
Giveup (18)
None (28)
Prompt (16)
Regular (28)
Condition
By Interaction



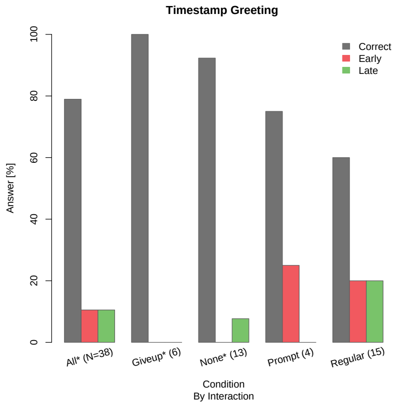
Timestamp
Answer [%]
0
20
40
60
80
100
After
Before
All (N=90)
Blocked (34)
Nopoint (28)
Unblocked (28)
Condition
By Blocktype



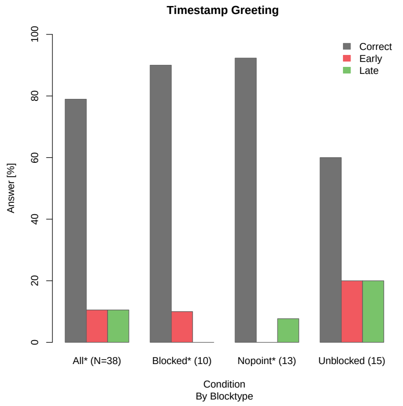
Timestamp Greeting
Answer [%]
0
20
40
60
80
100
Correct
Early
Late
All* (N=34)
Delayed (5)
Full* (19)
Random* (10)
Condition
By Opening



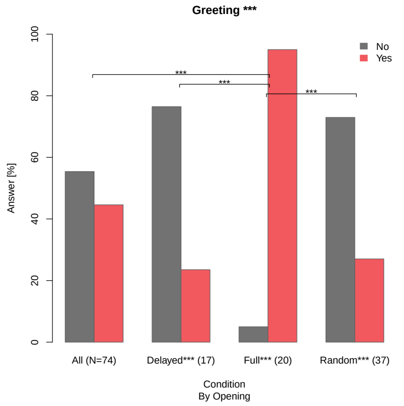
Timestamp Greeting
Answer [%]
0
20
40
60
80
100
Correct
Early
Late
All* (N=38)
Strong* (13)
Weak* (25)
Condition
By Entry



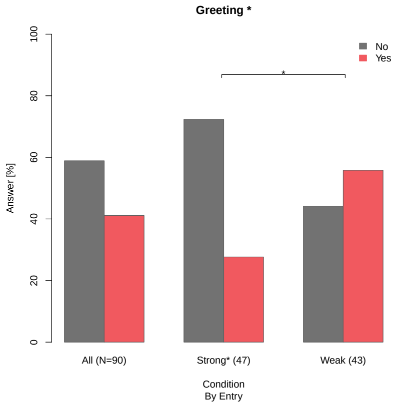
Timestamp Greeting
Answer [%]
0
20
40
60
80
100
Correct
Early
Late
All* (N=38)
Giveup* (6)
None* (13)
Prompt (4)
Regular (15)
Condition
By Interaction



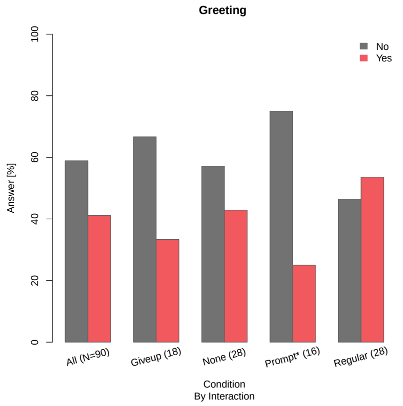
Timestamp Greeting
Answer [%]
0
20
40
60
80
100
Correct
Early
Late
All* (N=38)
Blocked* (10)
Nopoint* (13)
Unblocked (15)
Condition
By Blocktype



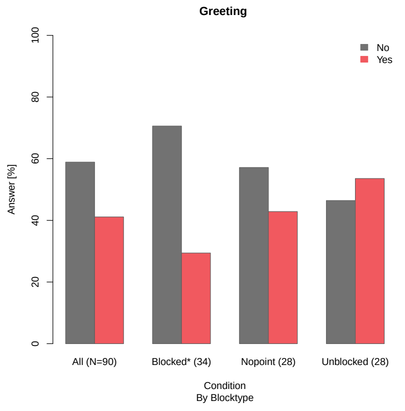
Greeting ***
Answer [%]
0
20
40
60
80
100
No
Yes
***
***
***
All (N=74)
Delayed*** (17)
Full*** (20)
Random*** (37)
Condition
By Opening

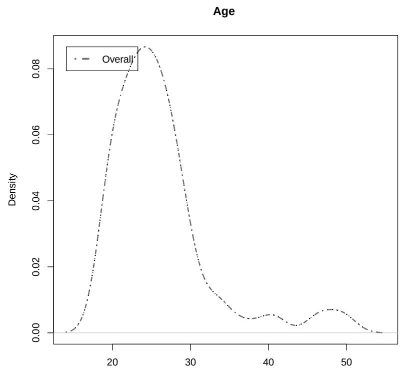


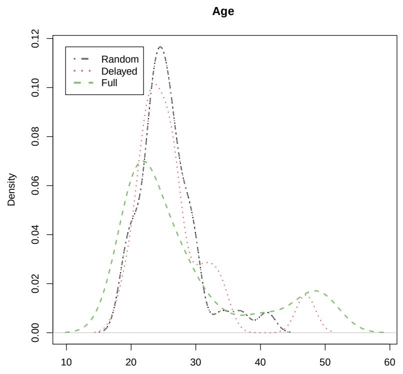
Greeting *
Answer [%]
0
20
40
60
80
100
No
Yes
*
All (N=90)
Strong* (47)
Weak (43)
Condition
By Entry

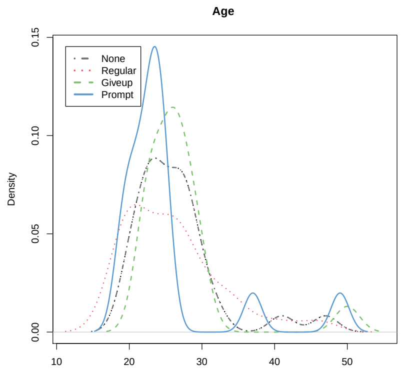


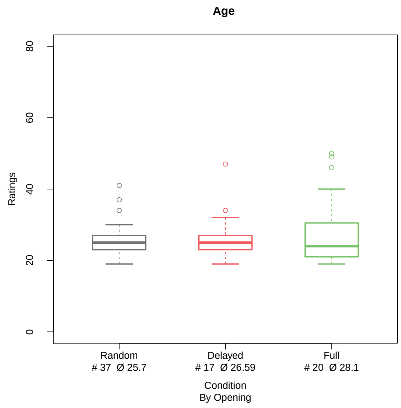
Greeting
Answer [%]
0
20
40
60
80
100
No
Yes
All (N=90)
Giveup (18)
None (28)
Prompt* (16)
Regular (28)
Condition
By Interaction

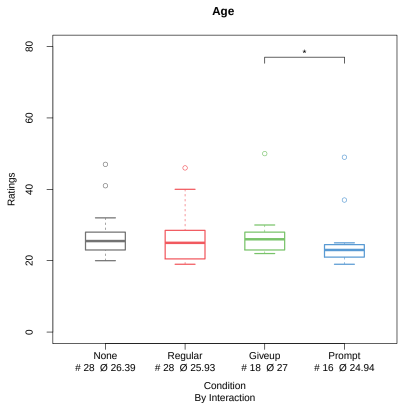

Greeting
Answer [%]
0
20
40
60
80
100
No
Yes
All (N=90)
Blocked* (34)
Nopoint (28)
Unblocked (28)
Condition
By Blocktype

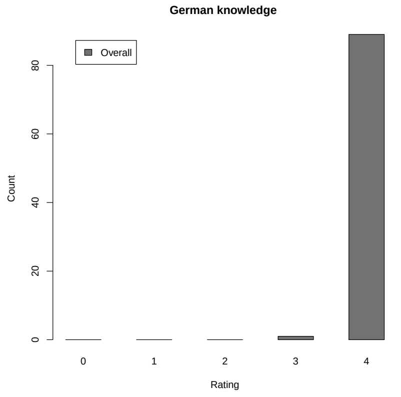

Age
Density
0.00
0.02
0.04
0.06
0.08
Overall
20
30
40
50

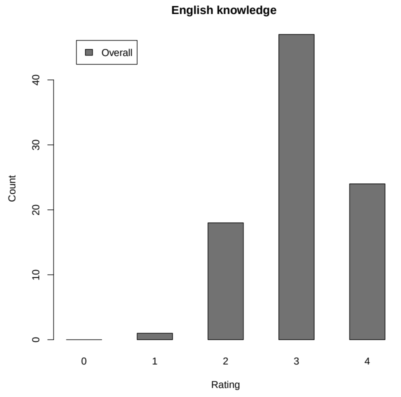

Age
Density
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Random
Delayed
Full
10
20
30
40
50
60

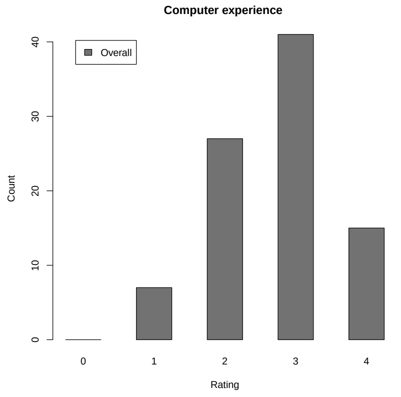

Age
Density
0.00
0.05
0.10
0.15
None
Regular
Giveup
Prompt
10
20
30
40
50

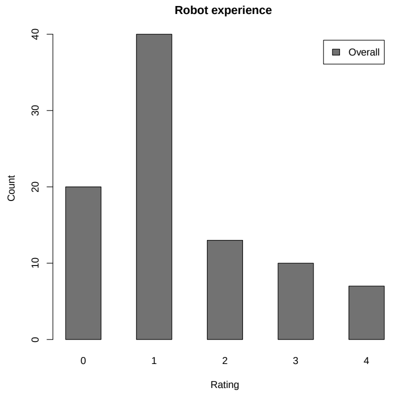

Age
Ratings
0
20
40
60
80
Random
37 Ø 25.7
Delayed
17 Ø 26.59
Full
20 Ø 28.1
Condition
By Opening

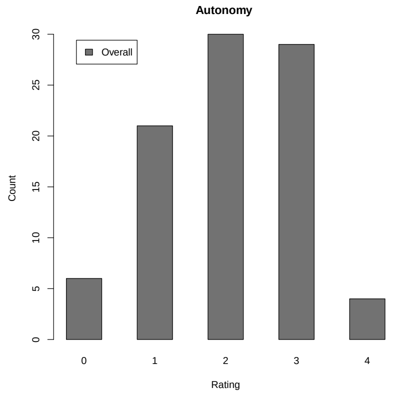

Age
*
Ratings
0
20
40
60
80
None
28 Ø 26.39
Regular
28 Ø 25.93
Giveup
18 Ø 27
Prompt
16 Ø 24.94
Condition
By Interaction

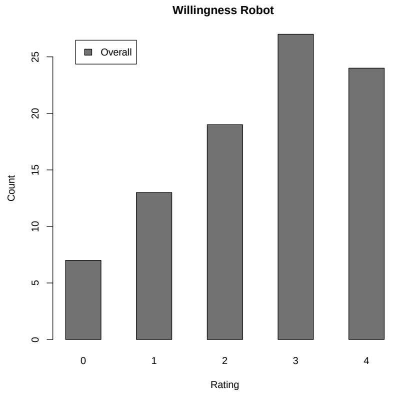

German knowledge
Count
0
20
40
60
80
Overall
0
1
2
3
4
Rating

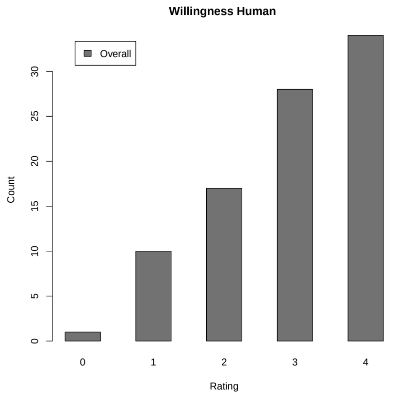

English knowledge
Count
0
10
20
30
40
Overall
0
1
2
3
4
Rating

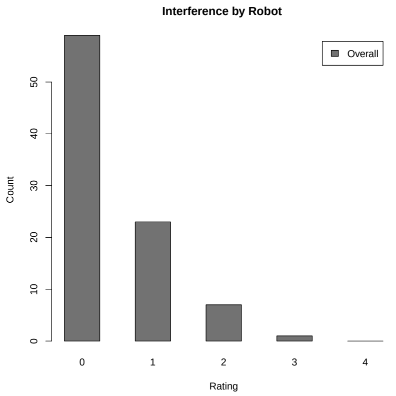

Computer experience
Count
0
10
20
30
40
Overall
0
1
2
3
4
Rating

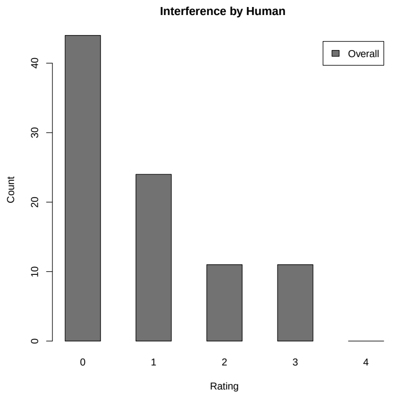

Robot experience
Count
0
10
20
30
40
Overall
0
1
2
3
4
Rating


Autonomy
Count
0
5
10
15
20
25
30
Overall
0
1
2
3
4
Rating


Willingness Robot
Count
0
5
10
15
20
25
Overall
0
1
2
3
4
Rating


Willingness Human
Overall
Count
0
5
10
15
20
25
30
0
1
2
3
4
Rating


Interference by Robot
Overall
Count
0
10
20
30
40
50
0
1
2
3
4
Rating


Interference by Human
Overall
Count
0
10
20
30
40
0
1
2
3
4
Rating