SOCIAL COGNITION AND ROBOTICS

Jasmin Isabella Kajopoulos

Graduate School of
Systemic Neurosciences
LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences der
Ludwig-Maximilians-Universität München

29.Mai 2019

Supervisor
Prof. Agnieszka Wykowska
IIT Central Research Laboratories
Istituto Italiano di Tecnologia (IIT)

First Reviewer:	Prof. Agnieszka Wykowska
Second Reviewer:	Prof. Hermann J. Müller
External Reviewer:	Prof. Patrick H. Khader

Date of Submission:	29.05.2019
Date of Defense:	05.08.2019

Table of Contents

I. List of Figures

II. List of Tables

Acknowledgement

I would like to thank my supervisor Prof. Agnieszka Wykowska for offering me the opportunity to conduct this Ph.D. thesis. Further, I am grateful to Prof. Hermann Müller and Prof. Gordon Cheng for extending their resources, supporting me and letting me conduct research in their laboratories. I also extend warm gratitude to Prof. Kazuhisa Takemura for providing me with the opportunity to conduct research at his laboratory. My thanks go also to my graduate school, the Graduate School of Systemic Neurosciences, who supported me, particularly in my first year and all the members of staff, who advised me.

My mother deserves special thanks for supporting me during this phase of my life. I am also grateful to my grandparents for supporting me and caring about my graduation.

Abstract

In our social world we continuously display nonverbal behavior during interaction. Particularly, when meeting for the first time we use these implicit signals to form judgments about each other, which is a cornerstone of cooperation and societal cohesion.

The aim of the studies presented here was to examine which gaze patterns as well as other types of nonverbal signals, such as facial expressions, gestures and kinesics are presented during interaction, which signals are preferred, and which signals we base our social judgment on. Furthermore, it was investigated whether cultural context, of German or Japanese culture, influences these interaction and decision making patterns.

One part of the following dissertation concerned itself mainly with gaze behavior as it is one of the most important tools humans use to function in the natural world. It allows monitoring the environment as well as signalling towards others. Thus, measuring whether attentional resources are captured by examining potential gaze following in reaction to pointing gestures and gaze shifts of an interaction partner was one of the goals of this dissertation. However, also intercultural differences in gaze reaction towards direct gaze during various types of interaction were examined.

For that purpose, a real-world dyadic interaction scenario in combination with a mobile eyetracker was used.

Evidence of gaze patterns suggested that independent of culture interactants seem to mostly ignore irrelevant directional cues and instead remain focused on the face of a conversation partner, at least while listening to said partner. This was a pattern also repeated when no displays of directional signals were performed. While speaking, on the other hand, interactants from Japan seem to change their behaviour, in contrast to interactants from Germany, as they avert their gaze away from the face, which may be attributed to cultural norms.

As correct assessment of another person is a critical skill for humans to possess the second part of the presented dissertation investigated on which basis humans make these social decisions. Specifically, nonverbal signals of trustworthiness and potential cooperativeness in Germany and in Japan were of interest.

Thus, in one study a mobile eyetracker was used to investigate intercultural differences in gaze patterns during the social judgment process of a small number of sequentially presented potential cooperation partner. In another study participants viewed video stimuli of faces, bodies and faces + bodies of potential cooperation partner to examine the basis of social decision making in more detail and also to explore a wider variety of nonverbal behaviours in a more controlled manner.

Results indicated that while judging presenters on trustworthiness based on displayed nonverbal cues German participants seem to partly look away from the face and examine the body. This is behavior in contrast to Japanese participants who seem to remain fixated mostly on the face. Furthermore, it was shown that body motion may be of particular importance for social judgment and that body motion of one’s own culture as opposed to a different culture seems to be preferred. Lastly, nonverbal signals as a basis of decision making were explored in more detail by examining the preferred interaction partner’s behaviour presented as video stimuli.

In recent years and presumably also in the future, the human social environment has been growing to include new types of interactants, such as robots. To therefore ensure a smooth interaction, robots need to be adjusted according to human social expectation, including their nonverbal behavior. That is one of the reasons why all results presented here were not only put in the context of human interaction and judgment, but also viewed in the context of human-robot interaction.

Introduction

Gaze Behaviour and Why It Fascinates Us

Gaze plays a special role in regard to social interaction, which is most likely owed to its dual nature. On the one hand, with our eyes we perceive the environment, which allows us to monitor our surroundings and it enables us, specifically during interaction, to receive signals from our interaction partner (Gobel, Kim, & Richardson, 2015; Jarick & Kingstone, 2015).
On the other hand, gaze itself, but also the region immediately surrounding the eye, can send an extensive amount of information to others (Frischen, Bayliss, & Tipper, 2007).

Morphological Uniqueness of Eyes

This type of nonverbal communication via gaze is facilitated by the unique morphological nature of the human eye. In contrast to other primates, humans possess a large white sclera surrounding a smaller, dark pupil (with iris). In that way it seems that while most mammals are adapted to camouflage, their gaze being hidden to escape predators, humans seem to expose their gaze direction, perhaps due to having adapted towards social interaction (Kobayashi & Kohshima, 2001). Thus, it seems that over the course of humans' evolutionary history gaze has evolved and gained, besides its perceptual capabilities, also a communicative meaning. This so-called "cooperative eye" hypothesis proposes that social gaze evolved as humans developed cooperative societal structures (Tomasello, Hare, Lehmann, & Call, 2007).

Another piece of evidence in favour of the morphological and – in consequence – also communicative uniqueness of the human eye has been provided by Emery (2000). He described how certain anatomical features of the face, for example, face protrusions, cheekbones, the shape of nose and eyebrows, have evolved in a way that the eyes became more clearly visible enabling humans to precisely determine direction of others' gaze, i.e. direction of attention (Emery, 2000).
This has been further investigated with studies on assorted eye stimuli. It has been demonstrated that either inverted sclera and pupil colour, i.e. dark sclera with white pupil (Ricciardelli, Baylis, & Driver, 2000; Sinha, 2000), or partially greyed sclera (Ando, 2004) or even masked eye region led to changed and often reduced gaze perception or impaired face detection (Lewis & Edmonds, 2003) and recognition (McKelvie, 1976).

Importance of Eye Region.

The eye region is key for several aspects of social interaction. It is central to categorizing gender (Schyns, Bonnar, & Gosselin, 2002), distinguishing identity (Vinette, Gosselin, & Schyns, 2004) and particularly important for recognizing certain emotional expressions, such as anger, sadness and fear (Calder, Young, Keane, & Dean, 2000). Similarly, eyes seem to be involved in expressing joy, and surprise (Ekman, Friesen, & Tomkins, 1971), but have also been shown to be particularly crucial for recognizing threat in others (Fox & Damjanovic, 2006) or distinguishing between real and fake smiles in others (Ekman, Davidson, & Friesen, 1990), and may also play an important part in discriminating more complex mental states, as has been shown with the relatively well-known “Reading the mind in the eyes” test (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).

Direct Gaze and Its Function

Direct gaze contact appears to be of particular significance to social interaction. At 5 months, infants have the ability to distinguish between horizontal gaze deviations of up to 5° (Symons, Hains, & Muir, 1998) and adult humans can discriminate between direct and averted gaze at already 2.8 ° (Gibson & Pick, 1963). Furthermore, direct gaze as opposed to averted gaze is associated with enhanced face processing and recognition, gender discrimination and has also been shown to affect social judgment of an interaction partner (Hood, Macrae, Cole-Davies, & Dias, 2003; Macrae, Hood, Milne, Rowe, & Mason, 2002; Mason, Hood, & Macrae, 2004). This preference for direct gaze has also been demonstrated in infants of 4 months (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Johnson, & Csibra, 2004). Although direct gaze may be interpreted as threatening, aggressive or dominant (Kleinke, 1986), depending on the social context direct gaze may also lead to increased attraction (Mason, Tatkow, & Macrae, 2005). Other associations with direct gaze have been increased credibility, competence and also greater perceived social skills (Kleinke, 1986). Based on these numerous implication of the profound social impact of gaze contact, it has been proposed that direct gaze signals may in general simply signify a subsequent positive or negative social interaction (Itier & Batty, 2009), which therefore seem to elicit emotional responses in the receiver (Nichols & Champness, 1971).

Deictic “Pointing” Gaze and Other Directional Cues

Averted gaze, on the other hand, signals that a person is attending to something else in the environment. Specifically, it reveals to the observer the visual focus and object of interest of the gazing interaction partner. By imagining or mimicking the observed behaviour, i.e. gaze following, attention is shifted to the same visual focus as the original gazer’s. Thus, based on experiencing the same point-of view as the interaction partner the observer may now infer the others internal mental states such as intentions, desires, goals and future actions. Therefore, gaze as a directional social cue is a key signal for triadic coordination between two individuals and an object, wherein both individuals jointly attend the object and each other. Thus, deictic gaze and gaze following evoke the sophisticated ability of theory of mind (Baron-Cohen, 1995; Mundy, 2017; Shepherd, 2010; Tomasello, 1995). As gaze following and attentional orienting elicited by directional social cues is fundamental to inform our impressions about the environment we share and also about the persons we interact with (Bayliss, Frischen, Fenske, & Tipper, 2007; Bayliss & Tipper, 2006), it has been investigated in abundance. Nevertheless, the question remains whether we have a bottom-up, reflexive tendency to follow observed deictic gaze or whether we voluntarily shift attention in a top-down controlled manner (Theeuwes, 2010). For example, preconceptions on mental states or action plans of an individual as well as previous gaze contact have been demonstrated to modulate attentional orienting (Kompatsiari, Ciardo, Tikhanoff, Metta, & Wykowska, 2018; Perez-Osorio, Müller, Wiese, & Wykowska, 2015; Wiese, Wykowska, Zwickel, & Müller, 2012).

Comparably to directional gaze also pointing gestures may be able to influence attentional orienting. In fact, pointing has been demonstrated to induce, similar to gaze cues, reflexive orienting towards a target (Hamilton, 2017; Langton & Bruce, 2000).

Importantly, one has to bear in mind that some nonverbal signals, particularly hand gestures and also gaze movements, may in real life interaction not necessarily retain any intentional communicative function, but may instead rather be employed unintentionally in support of speech formulation (Krauss, Chen, & Chawla, 1996).

Gaze Behaviour – Own Research

In summary, gaze behaviour is an essential part of social interaction, which is why I dedicated two chapters, Chapter 1 and Chapter 2, to its understanding. Chapter 1 concerns itself mainly with directional gaze. Most research in the field of directional gaze has been conducted by using isolated 2D stimuli, i.e., faces or cartoons. However, during natural interactions gaze occurs in combination with other social signals, for example hand gestures. It has also been proposed that instead of using static stimuli to examine social mechanisms, participants should instead be immersed in a real social interaction (Cole, Skarratt, & Kuhn, 2016; Schilbach et al., 2013; Skarratt, Cole, & Kuhn, 2012). To merge the need for real life interaction to elicit social cognitive responses as well as the need for combined social signals, I designed my experiments in such a way that a trained experimenter interacted with participants. In this manner, I could examine whether in a real-life interaction scenario directional gaze also in combination with directional gestures, would evoke gaze following in participants in a reflexive manner (Chapter 1). Chapter 2, on the other hand, deals with the study of direct gaze and investigates under which interaction conditions direct gaze may be elicited in a real interaction setting. One of the conditions I have been interested in here, was culture (see paragraph Culture and Gaze). Both Chapters 1 and 2 were based on research conducted by Kajopoulos (2015).

Culture and Gaze

Previous studies have argued that social signals, such as gaze or pointing, may be culturally determined, making it advisable, especially in the context of social interaction scenarios, to study

these signals across different cultures (Ellsworth & Ludwig, 1972; Kleinke, 1986).
Accordingly, LaFrance and colleagues (1976) investigated subcultural differences of gaze behaviour in the USA and found participants looking less at an interaction partner while listening than while speaking (LaFrance & Mayo, 1976, 1978), which is a reversal of the typical pattern, i.e. increased gaze activity towards the interaction partner while listening as compared to speaking (Argyle & Cook, 1976; Bavelas, Coates, & Johnson, 2002; Kendon, 1967) Similarly, differences in gaze behaviour based on explicit culture-specific norms and rules have been reported by Rossano (2013). In some cultures gaze contact during conversation is considered rude, for example in Navaho Indian culture, or even completely avoided, for example, cultural norms of the Luo people in Kenya prescribe that a man and his mother-in-law should be positioned back-to-back during conversation (Argyle & Cook, 1976). Similarly, people of indigenous Mayan cultural background appear to be relatively gaze aversive while communicating (Rossano, Brown, & Levinson, 2009) as compared to the typical behaviour observed of people from the "Western culture" of Italy, i.e. looking mostly at the face of an interaction partner while listening if no other task is at hand (Rossano, 2013).

As gaze behaviour may therefore to some extend depend on the culture within which a person develops, misinterpreted signals may lead to inter- and subcultural communication difficulties. (Ellsworth & Ludwig, 1972). To avoid these pitfalls, deciphering such differences in gaze patterns may be of high importance. Particularly, for cultures that value nonverbal abilities, and where they are necessary to understand the context of verbal communication. For example in Japanese culture, examining gaze behaviour may thus be of advantage. (McDaniel, 1993; Morsbach, 1973).
In this vein the so-called "social display rules hypothesis" has been proposed, which states that social and cultural norms may modulate otherwise endogenous gaze behaviour. Evidence in favour of this hypothesis has been reported by McCarthy and colleagues (2008), who demonstrated that Canadian participants maintained gaze while thinking about an answer only when they believed to be observed. In cases where they felt to be unobserved Canadian participants seemed to avert gaze. Japanese participants on the other hand, consistently showed to avert their gaze while thinking about an answer independent of their belief about being observed, which indicates that not only social but also cultural context influences gaze behaviour (McCarthy, Lee, Itakura, & Muir, 2008). Furthermore, it has been suggested that Japanese participants scan faces of anthropomorphic virtual characters differently than British participants. First, Japanese participants were shown to fixate the avatars' eyes more, whereas British participants fixated the avatar's mouth. Secondly, as soon as the avatar, whose head was always half-turned away from observers, moved its gaze further away, Japanese participants fixated longer on the more visible eye, which more clearly showed gaze direction, as if they were trying to find out in which direction the avatar looked, whereas British participants seemed more likely to ignore the gaze shift (Senju et al., 2013).

Interpreting the direction of gaze has also been shown to be modulated by culture. Uono and Hietanan (2015) reported that Japanese participants considered even slightly averted gaze to be eye contact independent of the gaze being displayed by faces of their own culture vs. a different culture. In contrast, this pattern was reported for Finish participants only for faces of the less familiar culture as compared to faces of their own culture (Uono & Hietanen, 2015).
In the context of all the evidence regarding cultural differences related to gaze behaviour, one needs to keep in mind that similarities of gaze patterns across Japanese and Western cultures have been found as well. While being interviewed, for example, also Japanese students, reflecting accounts of Western gaze behaviour, seem to respond with faster answers and less torso movement when being gazed at (Bond & Komai, 1976). Other studies have shown that gazing induces faster responses in multiple cultures, including various "Western" cultures. Although overall to a smaller degree a similar trend was found also for Japanese culture (Stivers et al., 2009).

Culture and Gaze – Own Research

As culture may be one of the fundamental factors that influence social interaction, all research presented in the following chapters (Chapter 1 - 4) was conducted in two different cultures, in Germany and in Japan. Particularly, research presented in Chapter 2 (based on Kajopoulos, 2015) investigated cultural influence on gaze behaviour during various types of interaction, such as listening, being asked a question and answering.

Trust and Trustworthiness

As humans live their life as part of various social groups, acts of cooperation and social exchange are common occurrences. One of the prerequisites for cooperation is trust, which has been defined as aa belief of common interests (Boone & Buck, 2003; Hardin, 2002). Although it has been claimed that trust can only be established after multiple "proven" positive actions of a long-term relationship, others have demonstrated the existence of trust between initial strangers (Berg, Dickhaut, & McCabe, 1995; Lount Jr, 2010). Trustworthiness, on the other hand, is seen as the perceived trait individuals are judged by that suggests security in the face of one's own vulnerabilities and investments (Boone & Buck, 2003; Hardin, 2002). Therefore, making accurate judgments of an interaction partner's trustworthiness is a critical skill for achieving successful social exchange. Similarly, being oneself perceived as trustworthy and therefore convincingly communicating trustworthiness is not only a necessity, but also an advantage from an evolutionary perspective (Boone & Buck, 2003). But what are the communicative methods with which we signal trustworthiness?

It has been argued that many signals of trustworthiness seem to be implicit and thus communicated nonverbally (Boone & Buck, 2003). This has been supported by ample evidence which shows that nonverbal communication alone is enough to engage in cooperation. Therein, accurately predicting cooperation or deceit in an economic game was based solely on a preceding 30-minute interaction without the interactants verbally communicating about the strategies they planned to employ in the following game (Kurzban, 2001).

Similarly, during counsellor – patient interactions judgments of trustworthiness from both sides seemed to mainly be based on nonverbal cues as opposed to communication content (Lee, McGill, & Uhlemann, 1988) and also in interview situations observers rated trustworthiness in light of the actions not the verbal cues (Kaul & Schmidt, 1971).

Furthermore, it has been observed that even non-human primates display behaviour that may be interpreted as signalling trustworthiness. In this context, dominant females used nonverbal signals to communicate to their less aggressive counterparts that they would desist from acting aggressively for the duration of a grooming episode. Hence, they seemed to signal guaranteed security towards the less dominant females and thus assured their trustworthiness to the benefit of a mutually cooperative action (Silk, 2002; Silk, Kaldor, & Boyd, 2000). Moreover, it has been shown that face-to-face interactions facilitate correct judgment of trustworthiness for a following cooperative exchange as opposed to interaction via chat program (DeSteno et al., 2012).

Based on this evidence it is plausible to assume that trustworthiness is communicated by implicit nonverbal behaviours. However, this still leaves open the question: What are these nonverbal signals that imply trustworthiness?

Signals of Trustworthiness

Body motion has been shown to be a contributor in that regard, wherein expansive and frequent arm movements seemed to negatively impact perceived trustworthiness and were instead associated with dominance displays (Koppensteiner, Stephan, & Jäschke, 2016).
In a systematic study on different categories of hand gestures it has been demonstrated that gesticulation that is related to speech (ex., pointing gestures or finger-counting movements) as opposed to gestures of self-touching or object-touching had a more positive effect on observers (Maricchiolo, Gnisci, Bonaiuto, & Ficca, 2009). This has also been reflected by results from DeSteno and colleagues (2012), who showed four prevalent behavioural cues of avoidance and untrustworthiness, hand touch, face touch, arms crossed and lean away, three of which seem to belong to the category self-touch.

On the other hand, facial expressions, particularly fleeting temporal subtleties, have also been known to influence perception of trustworthiness. Krumhuber et. al (2007) demonstrated how different smile dynamics, i.e. smiles with longer onset and offset as well as shorter apex duration, were perceived as more trustworthy and led to higher amount of cooperation.

Boone and Buck (2003) on the other hand argued that emotional expressivity is directly related to perceived trustworthiness. Particularly for first impressions nonverbally expressive persons have been judged in various contexts as more positive (Ambady & Rosenthal, 1993; Sabatelli & Rubin, 1986). They further explained the existence of good and bad senders of emotion, independent of the emotional valence, i.e. positive or negative emotions. Importantly, good senders accurately communicate their emotions, so that a receiving interaction partner may easily and correctly decipher the conveyed emotion. This is then of mutual benefit to both parties as cooperation can ensure. Additionally, they argued that nonverbal signals that communicate emotions are a continuous source of feedback to an interaction partner and that in an ongoing interaction an emotionally expressive person more easily demonstrates their motivation, but likewise with much more difficulty hides their deceptions. Specifically, they discussed, also in the context of previous research conducted by DePaulo (1992) that the only way for someone to ascertain if subsequent deception may occur or whether they will prove to be trustworthy is to base their judgment on emotional leakage, which is much more visible and thus easier to distinguish in emotionally expressive individuals (Boone & Buck, 2003).
Lucas et al. (2016) however suggested that their automatic tracking method (facial expression, head pose and gaze direction as well as word count) showed people using overall positivity (frequency, strength, duration of positive expressions) and level of contempt to judge trustworthiness during a negotiation task rather than emotional expressivity (average variation of anger, contempt, disgust, joy, fear, sadness and surprise).
This seems to be to the detriment of actual, objective trustworthiness, measured by number of lies told during the negotiation task, as measures for perceived trustworthiness weren't predictive of actual trustworthiness (Lucas, Stratou, Lieblich, & Gratch, 2016).
Yet, they didn't include other body language besides head pose in their analysis, which gives rise to the notion that gestures, for example, are essential to judging trustworthiness and are potentially also involved in emotional expressivity.

The question arises here of whether, as also proposed by Lucas and colleagues (2016), participants overly rely on facial expressions or whether and in what regard hand gestures may modulate the judgment of trustworthiness. Avierez (2012) investigated this conundrum and found that combining faces of highly positive facial expressions with negative body language led to judging the emotional expression as more negative. Similarly, negative faces combined with positive bodies led to negative emotional valence judgment (Aviezer, Trope, & Todorov, 2012). This again highlights not only the

importance of body language for the interpretation of emotional expressions, but also that facial expressions in isolation may lead to different results than combinations of face and body.
The question remains whether a similar outcome would develop when investigating nonverbal behaviour of trustworthiness.

For that purpose, we investigated in a series of experiments (Chapter 4) whether participants, who could choose their cooperation partner for a subsequent economic game, would decide on different interaction partners based on whether they had previously seen only the face of the potential interaction partner or the whole person (face and body).

Trustworthiness and Culture

Nevertheless, the point persists whether any results from above mentioned previous research, are based solely on American or other Western cultures or whether a much more reserved culture, such as Japanese culture (Matsumoto & Ekman, 1989), would for example display the same tendency of judging emotionally expressive people as more trustworthy.
Furthermore, it may still be the case that within a culture there exist more and less emotionally expressive people. Schug and Matsumoto (2010) conducted their study in Japan and showed that within the context of Japanese culture cooperators seemed to be more emotionally expressive than non-cooperators. However, it is interesting to note that although nonverbal signals seem to differ significantly from culture to culture (Kita, 2009) only few studies have concerned themselves with directly investigating these differences, particularly in the context of trust and cooperation.
In a similar vein, Jack and colleagues (2012) demonstrated the existence of systematic differences between facial expression signals of Western and Eastern cultures (Jack, Caldara, & Schyns, 2012). This is in direct contrast, but not necessarily contradicting, to studies which suggest a universality in the facial expression of emotion (Ekman & Friesen, 1971; Ekman, Sorenson, & Friesen, 1969). On the contrary, Ekman and Friesen (1969b) proposed the existence of pancultural aspects, (affect “programs”, i.e. the facial muscle movements) as well as culturally variable aspects (affect antecedents), that modulate the visible facial expression (Ekman & Friesen, 1969b). Similarly, Jack et al. (2012) showed that participants from an Eastern culture expected changes occurring in the eye region in accordance of their representation of different emotions, whereas for the Western culture the eyebrows and the mouth were the facial feature that seemed to reflect different emotional states. Most importantly, in the Eastern culture gaze shifts seemed to be much more pervasive in representing various emotions, which prompts the assumption that gaze direction is not only involved in joint attention and attentional guidance, but also reflects the emotional state of a person, at least in East Asian cultures. As there seems to be a difference in how emotions are facially

expressed in different cultures, facial expressions could thus be misread by people from outside the culture, which may then likewise lead to misjudgement of trustworthiness

Trustworthiness – Own Research

To examine these cultural differences, I have designed two studies that investigate nonverbal signals in the face of cooperation. In the study discussed in Chapter 3 participants could choose a cooperation partner for a subsequent economic game from a group of real people who gave a brief self-presentation about themselves. Importantly, participants couldn't hear the content of speech, as they were wearing earplugs and sound-attenuating head phones, thus they made their judgments based solely on nonverbal cues. As I conducted this study in Germany and in Japan, I wanted to compare cross-culturally which type of nonverbal behaviour captured most attention, facial expressions or body motions and gestures. To take a leaf out of the decision making and attention research and examine this in detail, I used a mobile eyetracking system which shed light on where (be it face or body) participants looked during interaction. Particularly, I was also interested in whether participants displayed a different type of gaze pattern while looking at a trustworthy person of their choice compared to a non-trustworthy person, i.e. not of their choice, and whether this varied in different countries, i.e. Germany and Japan.

In the second study discussed in Chapter 4, a study set consisting of three experiments conducted in Germany and in Japan respectively, I presented participants with videos of their potential interaction partner, either only their face (Experiment 1), only their body (Experiment 2) or both face and body (Experiment 3), from which they had to choose their future cooperation partner. To compare how much influence facial expression had on the perception of trustworthiness, I compared choice results from Experiment 1 – Face and Experiment 3 – Face and Body. Further, to investigate how culture influences trustworthiness perception of body motion I added videos from the respective other culture to Experiment 2, where the face wasn't visible, and participants could choose their interaction partner based solely on videos of their body language. Thus, I examined whether participants from Japan would choose a cooperation partner from their own culture or from another culture (Germany) and if the same was true for German participants. Lastly, I investigated for each experiment, whether of all presented interaction partners someone was more likely to be chosen than others and I broke down and summarized the probable nonverbal behaviour this decision may have been based on.

Summary of Research Questions

In summary the following main questions were examined in this body of work.

Question1: Does attentional capture occur from directional nonverbal signals during real life interaction? (Chapter 1)

Question 2: Are there intercultural differences of gaze patterns during various interaction types? (Chapter 2)

Question 3: Are there intercultural differences of gaze patterns when making judgment of trustworthiness and choosing a cooperation partner? (Chapter 3)

Question 4: Are there different gaze patterns for people one trusts and prefers as cooperation partner as opposed to people one doesn't trust? (Chapter 3)

Question 5: How important is facial expression as compared to other body language when deciding on a cooperation partner based on trustworthiness in Japan and in Germany? (Chapter 4)

Question 6: Does a bias for in-culture body language exist as opposed to out-culture when making judgments of trustworthiness and deciding on a cooperation partner? (Chapter 4)

Question 7: What are the probable nonverbal signals of trustworthiness in Japan and in Germany? (Chapter 4)

Outlook for Robotics

Nowadays, with enormous technological advances we need to face inclusion in our social sphere of not only other humans but also machines. This has given rise to a new field of research, the human-computer- or human-robot-interaction (HRI) research. Although, in some regards computers seem to be intrinsically accepted as social actors (Nass, Steuer, & Tauber, 1994) continuous improvement on HRI is one of the aims of the involved scientific community. Particularly, in light of the increasing number of robotic applications and implementations of artificial intelligence into the social environment, for example solutions for elderly care or therapy for children diagnosed with autism-spectrum disorder (ASD) (Broekens, Heerink, & Rosendal, 2009; Cabibihan, Javed, Ang, & Aljunied, 2013; Kajopoulos et al., 2015).

One of the most prevalent features that are highlighted seems to be trust; trust in the technology, but also trust in computers and robots as social agents and interaction partners (Cassell & Bickmore, 2000; Van Mulken, André, & Müller, 1999). Thus, the next question is, how does one establish trust in a robot? Interestingly, there may exist cultural differences not only in behaviour towards humans, but accordingly also towards machines seen as social entities. Katagiri and colleagues (2001) for example showed that students from the US and Japan displayed different social behaviours towards helpful as compared to unhelpful computers (Katagiri, Nass, & Takeuchi, 2001). In the last section of this manuscript, Chapter 5, I not only summarized the results of my research on gaze behaviour and other nonverbal features displayed during human interaction and regarding trustworthiness attribution, but I also discussed these in the context of implementing them into robotics. For that purpose, I first outlined prevalent research in human-robot interaction.

Thus, the last question to be answered in this manuscript is:

Question 8: How does my research contribute to improvement of human-robot interaction?

Chapter 1 Attentional Capture

1.1 Introduction

Research Question 1:

Do people orient their attention towards irrelevant to the content of conversation directional cues (pointing, pointing + gaze) in a real-life dyadic interaction scenario?

Result 1:

No, they seem to remain fixated on the face and ignore the directional cues.
However, they may orient more towards the pointing location if the pointing gesture is accompanied by gaze and the pointed at location is close by.

Research Question 2:

Is there a difference on this effect of attentional orienting based on culture (Germany or Japan)?

Result 2:

No, participants of both cultures seem to equally remain fixated on the face and ignore the directional cues.

Gaze and pointing signals (Hamilton, 2017; Langton & Bruce, 2000) are deictic nonverbal cues, which are frequently deployed during social interaction. On the one hand, these signals can be used to glean an interaction partner's focus of attention and thus make inferences also about their intentions, desires and action plans. Furthermore, they are critical for evoking engagement in joint attention (JA). JA occurs during a triadic coordination between two individuals and an object: the two individuals attend to each other as well as to the object, thus sharing attention (Baron-Cohen, 1995; Mundy, 2017; Tomasello, 1995). On the other hand, during social interaction some nonverbal signals, such as hand movements, have been known to solely help in speech formation and hence have no communicative function per se (Krauss et al., 1996). Even shifting gaze, for example looking down, has been known to facilitate other cognitive functions, i.e. facilitate thinking behaviour, or be part of emotional expression, instead of signalling attention (Jack et al., 2012; Krauss et al., 1996; McCarthy et al., 2008).

Typical paradigms to examine social directional cues and subsequent attentional orienting in laboratory conditions are cuing paradigms, particularly gaze-cueing paradigms. In this type of paradigm, a picture or drawing of a face is presented centrally to the observers; after a while, the depicted face shifts the eyes towards a lateral location on the screen, whereupon a target stimulus is presented either at the indicated location or some other location. Better performance is usually observed when the target appears at the gazed-at (cued) location, as compared to other (uncued) locations.

It has been postulated that this so-called *gaze-cueing effect* (GCE) is based on a reflexive orienting mechanism (Driver et al., 1999; Friesen & Kingstone, 1998). However, there is evidence that besides the bottom-up, reflexive component of GCE, top-down factors may also influence attentional orienting in response to gaze cues. In that respect, attentional orienting was shown to be modulated by prior beliefs about the mental state(s) or inferred action plans of the observed interaction partner (Perez-Osorio et al., 2015; Teufel, Fletcher, & Davis, 2010; Wiese et al., 2012; Wykowska, Wiese, Prosser, & Müller, 2014).

However, in context of the above discussed function of deictic gaze and hand movements limitations of these paradigms become obvious. First, in most paradigms only one cue is presented to the observer, whereas during natural social interaction gaze shifts hardly ever occur in isolation, but are instead embedded in a multitude of other nonverbal signals and contexts. Secondly and most importantly, during most social cuing paradigms deictic gaze is mostly assumed to be an attention directing cue (either correct or incorrect, but nevertheless still directing) towards a target and not perhaps a modulator of conversation, part of emotional expression formation or irrelevant to the interaction. It is important to note that assumed directional cues may be meaningless or incongruent to the content of an interaction. For example, in a conversation a speaker might say: “I took a plane to go to Japan” and point or look to the right, which does not literally indicate the location of Japan. The conversation partner may follow the pointing and gaze shift, as it is a cue that according to well-known studies reflexively triggers attentional orienting (Friesen & Kingstone, 1998); however, based on the content of the conversation, it is clear that the gesture conveys no meaning and that there is no specific target to which the speaker is pointing. In essence, this is an irrelevant cue – a *distractor* even – as it may also be detrimental to orient towards the signal direction. Thus, the present study was designed to investigate the question whether such potentially attention-summoning but nevertheless incongruent and thus potentially attention-distracting signals would influence the allocation of attention in a naturalistic interactive setting.

1.1.1 Aim of Study

In more detail, the present study, which was based on an experiment of Kajopoulos (2015), investigated how gesture and gaze movements affect attentional capture in a situation where these seemingly directional cues are entirely ambiguous with respect to the conversation topic, that is, when they are potentially distracting. In other words: how reflexive, is the “*attentional-capture*” mechanism underlying orienting to directional social cues in naturalistic social scenarios?
For that purpose, the study eclipsed artificial laboratory stimuli and made use of a real-life

interaction partner that acted according to a predefined script. This interaction partner, henceforth called the experimenter, displayed natural, but scripted directional gestures and gaze shifts during a dyadic conversation with participants. Importantly, the gestures and gaze shifts were irrelevant to the content of conversation. The hypothesis on orienting mechanisms being reflexively employed by directional social cues (Friesen & Kingstone, 1998) as well as being the most salient cues in the environment (Theeuwes, 1994, 2004) would imply that participants would follow and direct their gaze towards the location of the deictic cues of the experimenter. On the other hand, should participants remain fixated on the face of the interaction partner and ignore the gaze shift and pointing gestures, attentional orienting may be more likely to be top-down controlled, as hypothesized by previous research on the modulability of attentional capture (Bacon & Egeth, 1994; Müller, Geyer, Zehetleitner, & Krummenacher, 2009; Wykowska & Schubö, 2010; Wykowska & Schubö, 2011). Thus, attentional capture was tested by the extent of participants' gaze direction, i.e. saccadic activity, following the observed deictic cues.

Importantly, in accordance with the hypothesis that social signals are not only embedded in but also regulated by the cultural context (Ellsworth & Ludwig, 1972; Kleinke, 1986) I conducted the study in a western culture – in Germany – as well as an eastern culture – in Japan.

Thus, the problem of a "Western" bias that has often plagued psychological research (Ellsworth & Ludwig, 1972; Kleinke, 1986; Pedersen, 2003) may be avoided. On the contrary, the present study may determine the cross-cultural differences or similarities of attentional capture effects in a real-life scenario.

1.2 Materials and Methods

1.2.1 Participants

25 German participants took part in the study at the Technical University Munich (TUM), Germany (Mean age: 25.16, SD = 4.99, 6 male). 5 participants were excluded as the eyetracker failed during calibration or data recording.
26 Japanese participants took part in the study at the Osaka Prefecture University (OPU), Japan (Mean age: 20.92, SD = 1.52, 12 male). 4 participants were excluded as the eyetracker failed during calibration or data recording and 2 participants were excluded as they indicated that they weren't naive to the experimental design.

During both experiments, in Germany and Japan, purely behavioural data, i.e., gaze activity and sound-recording of voices, was collected, and stored as well as analysed anonymously. Prior to the commencement of all experiments, participants gave their written informed consent. No invasive or potentially dangerous methods were involved, and all experimental procedures were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). For the participation in this experiment participants received either monetary compensation or course credit.

1.2.2 Apparatus and Experimental Setup

Gaze recording was carried out via mobile eye-tracker (SMI ETG, SensoMotoric Instruments GmbH, Germany, Fig. 1) with a scene camera resolution of 30 Hz, a gaze camera resolution of 60Hz and a positional accuracy of 0.5°. An integrated microphone recorded the sound. Experimenter and participant were seated opposite each other at a table at approximately 1m distance. While talking experimenters' hands were in the home position: visible on the table on top of a white piece of paper, holding a pen with both hands. While pointing the experimenter held the pen in the pointing hand (Fig. 2).

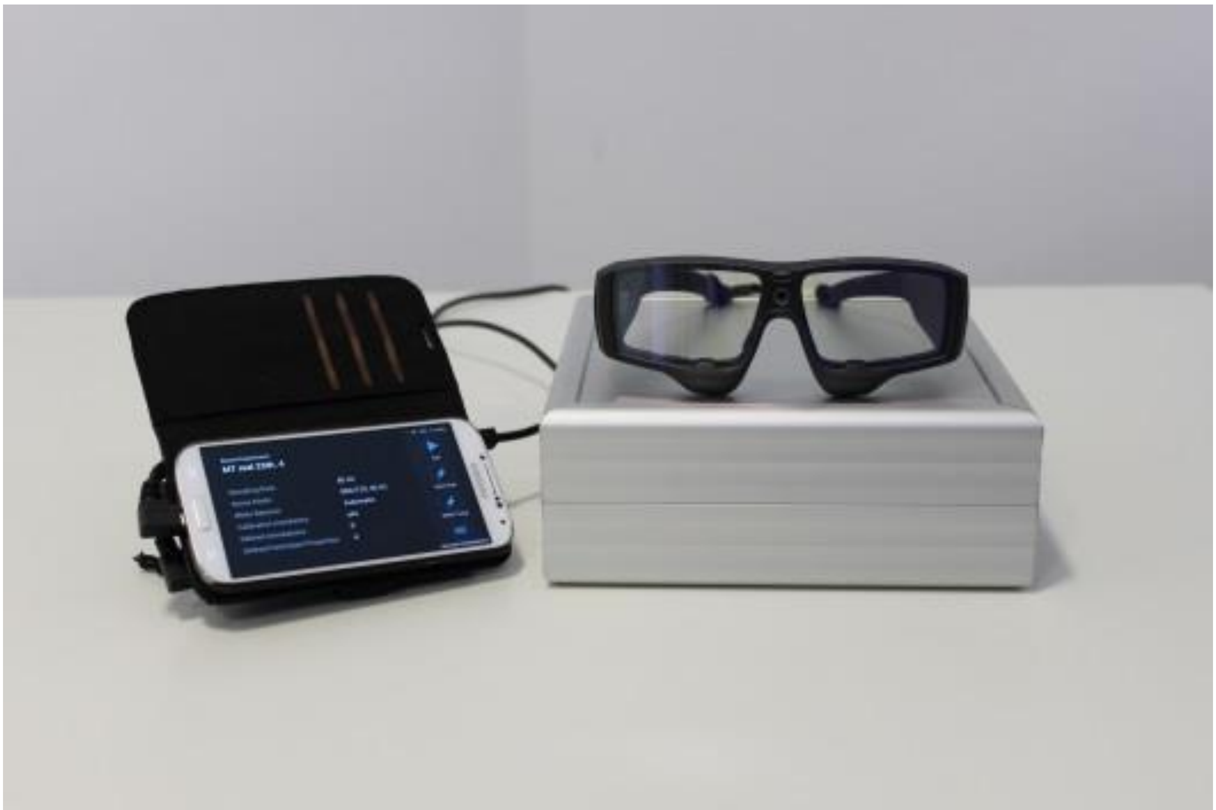

Fig. 1: Mobile eyetracker for tracking of gaze movements (SMI ETG, SensoMotoric Instruments GmbH, Germany)(from Kajopoulos, 2015).

Fig. 2: Experimental Setup. Dyadic interaction between the experimenter (left) who had memorized an interaction script and a participant (right) wearing the mobile eye-tracker whose gaze patterns were recorded (from Kajopoulos, 2015).

1.2.3 Procedure

This cross-cultural study comprised two experiments, one conducted in Germany with a native German speaking experimenter and one in Japan with a native Japanese speaking experimenter. Both experimenters wore neutral clothes, a black shirt and blue jeans. According to the interaction script, as a default, experimenters mostly showed a neutral facial expression, looked straight ahead towards the participant, and kept their hands in the home position, i.e. in front of them. However, during the critical gaze shift and pointing trials experimenters were explicitly instructed to act accordingly, i.e. move their gaze and/ or point.

1.2.3.1 Calibration

Preceding the gaze recording, 1-point calibration commenced. In rare cases when 1-point calibration failed, 3-point calibration was carried out instead. Afterwards, the gaze recording was switched on covertly and participants were asked to look at various fixation points on the body of the experimenter, which could be used in a second, offline calibration process.

1.2.3.2 Cover story

As the gaze recording was turned on covertly, participants weren't immediately aware of it. After the actual calibration participants were instructed that the calibration process was still ongoing automatically and that the eye-tracker needed time to adjust to their eyes. Further, they were instructed that the aim of the experiment was examining gaze patterns during reading for which purpose four questionnaires should be filled out after the calibration process finished. In the meantime, participants were told that while waiting the experimenter would like to hold an informal conversation with them. By making the participants believe the critical gaze-tracking period was the later held reading of the questionnaires instead of the current interaction, it was assured that participants' gaze behaviour during the interaction was as natural as possible. After the experiment participants were fully briefed about the cover story (Fig. 3).

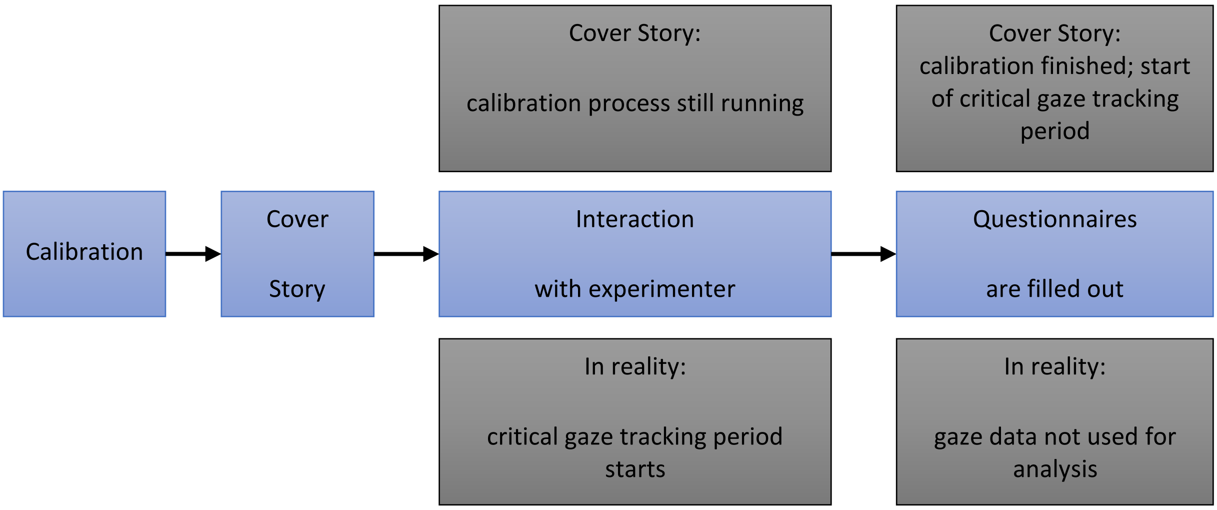

Fig. 3: Experimental procedure including the calibration process, cover story and critical gaze tracking period.

1.2.3.3 Interaction Script

During the period of the natural conversation with the participants, the experimenter acted according to a memorized interaction script.

Based on this script participants first listened to the experimenter talk, were subsequently asked a question, which they had then time to answer. This not only familiarized participants with the situation before the critical pointing phase but was also part of a bigger experimental study on general gaze behaviour during various types of interaction (see Chapter 2). This general phase was repeated 5 times [L1, Q1, A1; L2, Q2, A2; L3, Q3, A3; L4, Q4, A4; L5, Q5, A5; …].

This scripted rhythm of the interaction, listening (L)- being questioned (Q)- answering (A), continued also in the next most critical phase of the experiment. However, the one difference was that the experimenter demonstrated one of four types of pointing movements to non-meaningful locations during the listening phase of the interaction, i.e. while the participants listened and the experimenter talked. The first type of pointing movement shown was the pointing near movement (PN = listening phase + pointing near) after A5 and A6 (Fig. 4B). Next, the experimenter pointed near and gazed near at the same time (PNg = listening phase + pointing near + gaze) after A7 and A8 (Fig. 4D). Following A9 and A10 the experimenter showed far pointing movements (PF = listening phase + pointing far) (Fig. 4A). And lastly, after A11 and A12 the experimenter demonstrated far pointing movements while also gazing far (PFg = listening phase + pointing far + gaze) (Fig. 4C). [...A5, PN, Q6, A6, PN, Q7, A7, PNg, Q8, A8, PNg, Q9, A9, PF, Q10, A10, PF, Q11, A11, PFg, Q12, A12, PFg, Q13, A13...].

It was revealed that in most cases (Germany: 15, Japan:19) the experimenter followed the script exactly. Only in case the participant jumped ahead and already answered a question this question was skipped (Germany: 4). Whereas the order of scripted trials was only slightly altered in few cases

(Germany: 1 (PFg, Q12, A12 exchanged with PF, Q11, A11.), Japan: 1 (... PNg, Q8, A8, PF, Q11, A11, PNg, Q9, A9, PF, Q10, A10 ...))

Lastly the interaction script continued with a Question and Answer phase until A25. Although, irrelevant to the attentional capture research question this was part of a bigger experimental study on general gaze behaviour during various types of interaction (see Chapter 2). [Q14, A14, Q15, A15, Q16, A16, Q17, A17, Q18, A18, Q19, A19, Q20, A20, Q21, A21, Q22, A22, Q23, A23, Q24, A24, Q25, A 25].

The following number of directional bids (pointing, pointing + gaze) were on average performed by the experimenter:

- pointing near (PN): 10.92 (SD = 2.54) events;
- pointing near + gaze (PNg): 11.00 (SD = 2.14) events;
- pointing far (PF): 12.77 (SD = 3.24) events;
- pointing far + gaze (PFg): 11.25 (SD = 1.67) events;

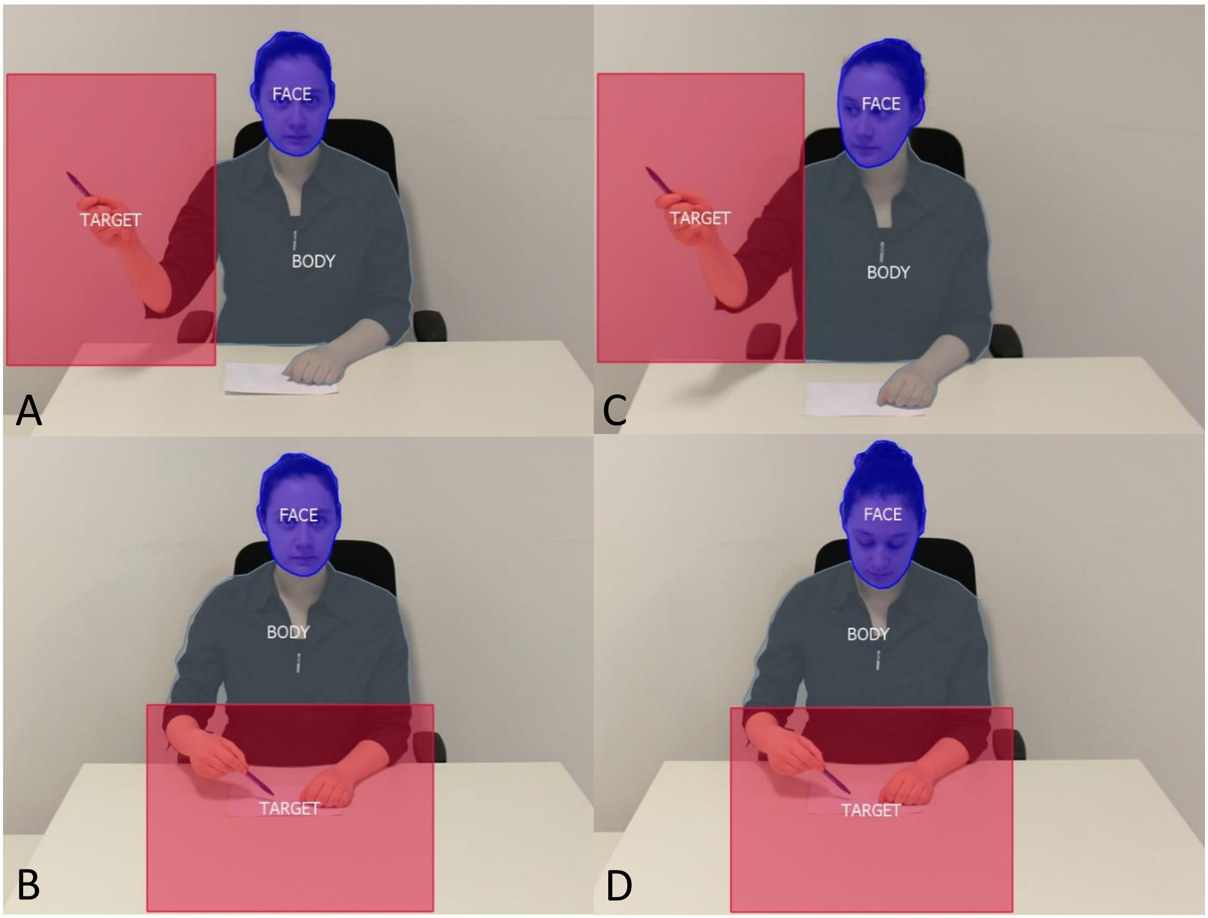

Fig. 4: Experimental Procedure with Pointing Phases and 4 Areas of Interest (Face, Body, Target, White Space). A= pointing far (PF); B = pointing near (PN); C = pointing far + gaze (PFg); D = pointing near + gaze (PNg). Areas of Interest: Face (blue), Body (grey), Target (red), everything else is Area of Interest White Space (adapted from Kajopoulos, 2015).

1.2.4 Data Analysis

The BeGaze software (SensoMotoric Instruments GmbH, Germany) was used to extract and label fixation events based on which Area of Interest they fell into (see Fig.4).

Only fixation events after onset of pointing movement until the pointing hand returned into the home position were included in the analysis. Importantly, as a mobile eye-tracker was used to record gaze movements the head of participants was constantly in motion. Also, the stimulus, i.e. the experimenter displaying pointing bids, was moving. This affected Areas of Interest (AOIs) who had to be defined based on dynamic real-world landmarks instead of fixed camera coordinates (pixel size or square centimetres). Thus, a reference image of the real world was used to label every gaze point with a particular AOI.

Specifically, if fixations fell upon the experimenter's pointing hand or in the direction of the pointing movement the label Area of Interest – Target was used. The Target AOI included approximately 11.5% of the whole reference. In cases where fixations fell on the face or upper body of the experimenter (excluding the Target AOI) the label Area of Interest (AOI) – Face and Body were used, respectively. All other fixations were labelled with White Space.

1.2.4.1 Fixation Time [%] Analysis

For every participant the fixation times (sum of durations of all fixations) were calculated relative to AOI, Pointing Distance (Far, Near), and Gaze condition (With Gaze, Without Gaze). Afterwards, mean Fixation Time [%], i.e., mean fixation time / epoch duration (PN, PNg, PF, PFg) was calculated.

The following average epoch durations were extracted from BeGaze:

- PN: 14.27s (SD= 5.52); PNg: 12.03 s (SD = 2.97);
- PF trials, 13.27s (SD = 3.88); PFg trials, 10.86s (SD = 3.15);

A mixed-design Analysis of Variance (ANOVA) with Distance (Far, Near), Gaze (With Gaze, Without Gaze), and AOI (Face, Body, Target, White Space) as within-subjects' factors and culture (Germany, Japan) as between-subjects factor was performed on Mean Fixation Time [%].

Mauchly's test for sphericity was generated and if significant the Greenhouse-Geisser corrected p-values *p[GG]* were used.

95% confidence intervals were calculated based on the correction method from Morey (2008), which takes within-subjects factors into account.

1.2.4.2 Scanpath Analysis

EyePatterns (Rochester Institute of Technology, 2006, Version 0.9, USA), a scanpath sequence analysis tool developed by West et al. (2006), was used to calculate transitions between subsequent AOIs. Markov Models are the basis of the transition matrix calculation (Eraslan, Yesilada, & Harper, 2015; West, Haake, Rozanski, & Karn, 2006). Specifically, every row (i) and column (j) in the 4x4 transition matrix denotes one of the 4 AOIs (Face, Body, Target, White Space) the gaze points in the sequence could fall upon. Every cell in the matrix has 3 entries, one entry for number of tallies between AOI i and AOI j, one entry for probability of AOI j following AOI i (%) and the third entry for probability of AOI i preceding AOI j (%C).

First a full-sequence transition matrix for all gaze points in the sequence was calculated. Secondly, a collapsed transition matrix, which included in the sequence only gaze points based on transitions between AOIs but not within AOIs, was calculated.

1.2.4.3 Probability of Fixations Analysis

To investigate the sequence of gaze events, particularly the first four fixation events after onset of directional stimuli in more detail, the average probability of fixations falling upon a specific AOI was analysed for a combination of every Distance (Far, Near) and Gaze (With, Without Gaze) condition.

In detail, the probability of fixation, i.e. number of participants whose fixations fell on particular AOI/ number of fixations across all AOIs, was calculated for any of the first four fixations.

1.3 Results

1.3.1 Analysis of Variance- Fixation Time %

The mixed-design ANOVA revealed a significant main effect of Area of Interest, $F(3, 114) = 335.94$, $\eta^2 = 0.86$, $p[GG] < .001$.

Post hoc paired t-tests (Bonferroni- corrected for multiple comparisons) indicated that the Face region was fixated more than the Body, Target, or White Space regions. However, comparing fixations on the Body, Target and White Space regions with each other resulted in similar amounts of gaze activity being found, all $p > .37$.

- Face – Body: $t(39) = 21.73$, $p < .001$
- Face – Target: $t(39) = 17.05$, $p < .001$
- Face – White Space: $t(39) = 24.84$, $p < .001$
  - Face: Mean (M) = 68.68, 95% confidence interval (CI) [63.38, 73.98]
  - Body: M = 4.33, 95% CI [2.01, 6.65]
  - Target: M = 4.53, 95% CI [0.60, 8.46]
  - White Space: M = 2.17, 95% CI [0.45, 3.89]

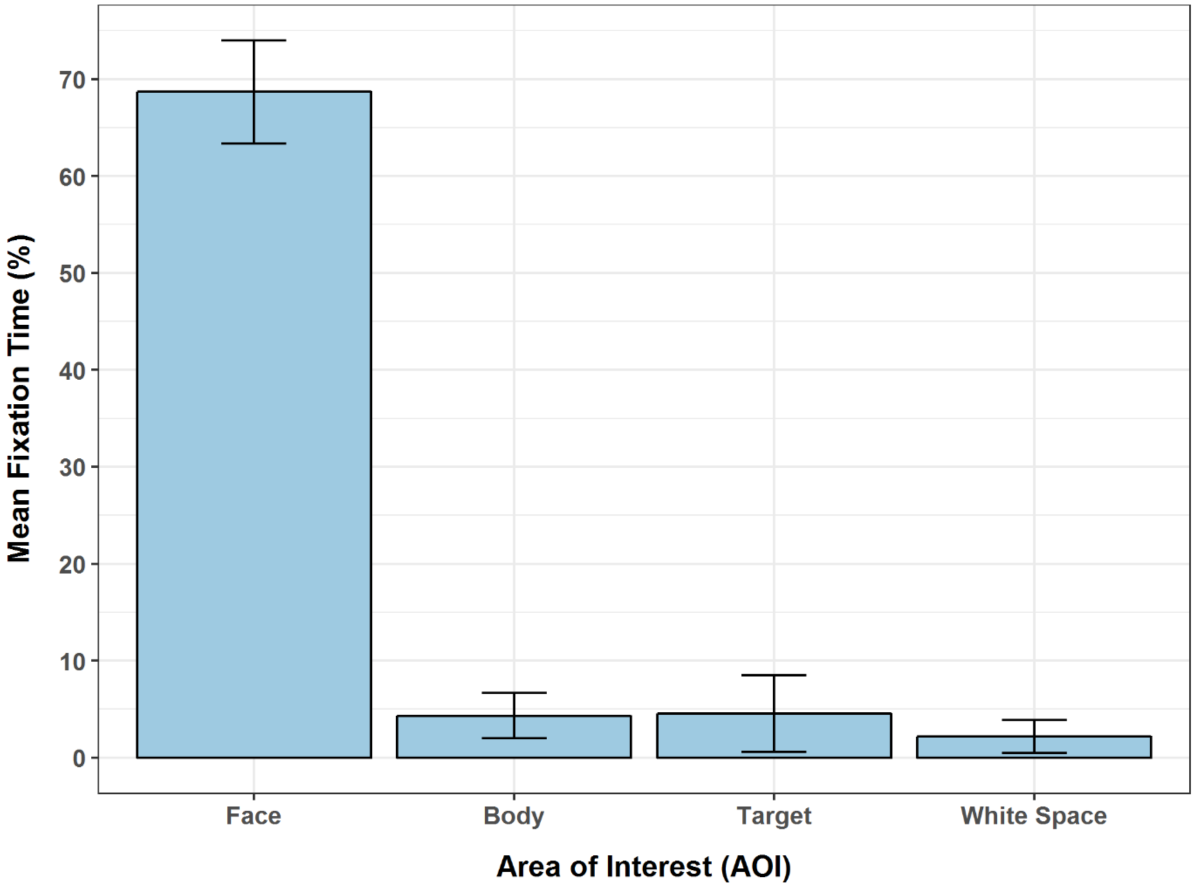

Fig. 5: Mean Fixation Time (%) for Area of Interest (Face, Body, Target, White Space); Error bars: 95% confidence interval (CI).

Furthermore, results yielded a significant interaction effect of AOI x Gaze, $F(3, 114) = 6.45$, $\eta^2 = 0.01$, $p[GG] = .006$ (Fig. 6).
Post-hoc comparisons revealed that significantly more fixation activity was centred on the Body when pointing occurred With Gaze than Without Gaze, $t(39) = 3.78$, $p = .002$.

- Body x With Gaze: M = 5.55, 95% CI [3.36, 7.74]
- Body x Without Gaze: M = 3.11, 95% CI [0.79, 5.43]

However, similar amounts of fixation activity were centred on the Face, Target and White Space, for pointing With Gaze and pointing Without Gaze, all $p > .09$ (Fig. 6).

- Face x With Gaze: M = 66.68, 95% CI [61.57, 71.79]
- Face x Without Gaze: M = 70.67, 95% CI [65.31, 76.03]
- Target x With Gaze: M = 5.58, 95% CI [1.61, 9.55]
- Target x Without Gaze: M = 3.49, 95% CI [< .001, 7,07]
- White Space x With Gaze: M = 2.25, 95% CI [0.59, 3.91]
- White Space x Without Gaze: M = 2.09, 95% CI [0.45, 3.73]

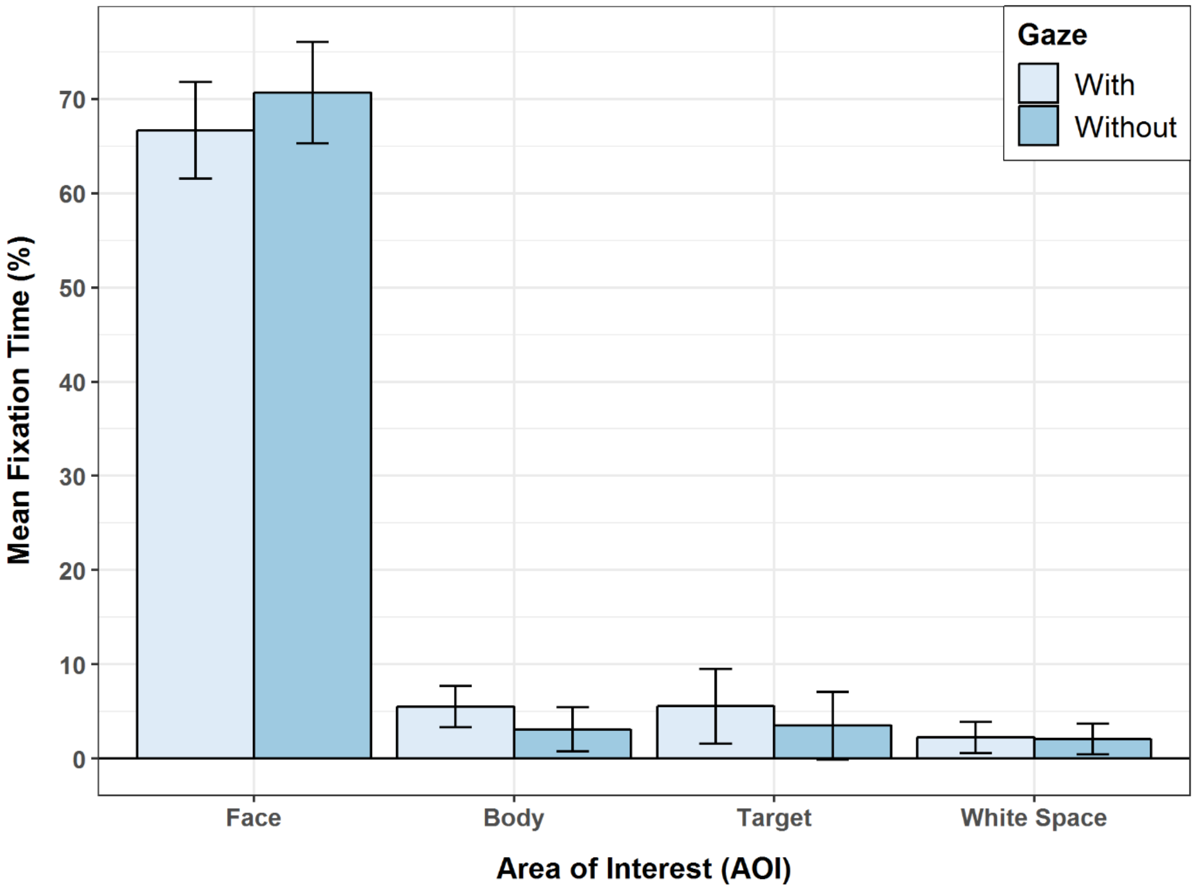

Fig. 6: Mean Fixation Time (%) for Gaze (With Gaze, Without Gaze) and Area of Interest (Face, Body, Target, White Space); Error bars: 95% CI.

Also, the AOI x Pointing Distance interaction was found to be significant, F(3, 114) = 8.26, = 0.02, *p[GG]* = .001 (Fig. 7).
Post-hoc comparisons showed that significantly more fixation activity was centred on the Face for the Pointing Distance Far than for the Pointing Distance Near, t(39) =3.07, p = .02.

- Face x Far: M = 71.56, 95% CI [66.15, 76.97]
- Face x Near: M = 65.79, 95% CI [60.58, 71.0]

In contrast, significantly less fixation activity was centred on the Body for the Pointing Distance Far than for the Pointing Distance Near, t(39) = -3.32, p = .008.

- Body x Far: M = 2.47, 95% CI [0.47, 4.47]
- Body x Near: M = 6.19, 95% CI [3.33, 9.05]

However, similar amounts of fixation activity were centred on the Target as well as the White Space, whether the Pointing Distance was Far (Target M = 3.82, 95% CI [0.50, 7.14]; White Space M = 2.04, 95% CI [0.53, 3.55]) or Near (Target M = 5.24, 95% CI [1.0, 9.48]; White Space M = 2.30, 95% CI [0.50, 4.10]), all *p* > .71 (Fig. 7).

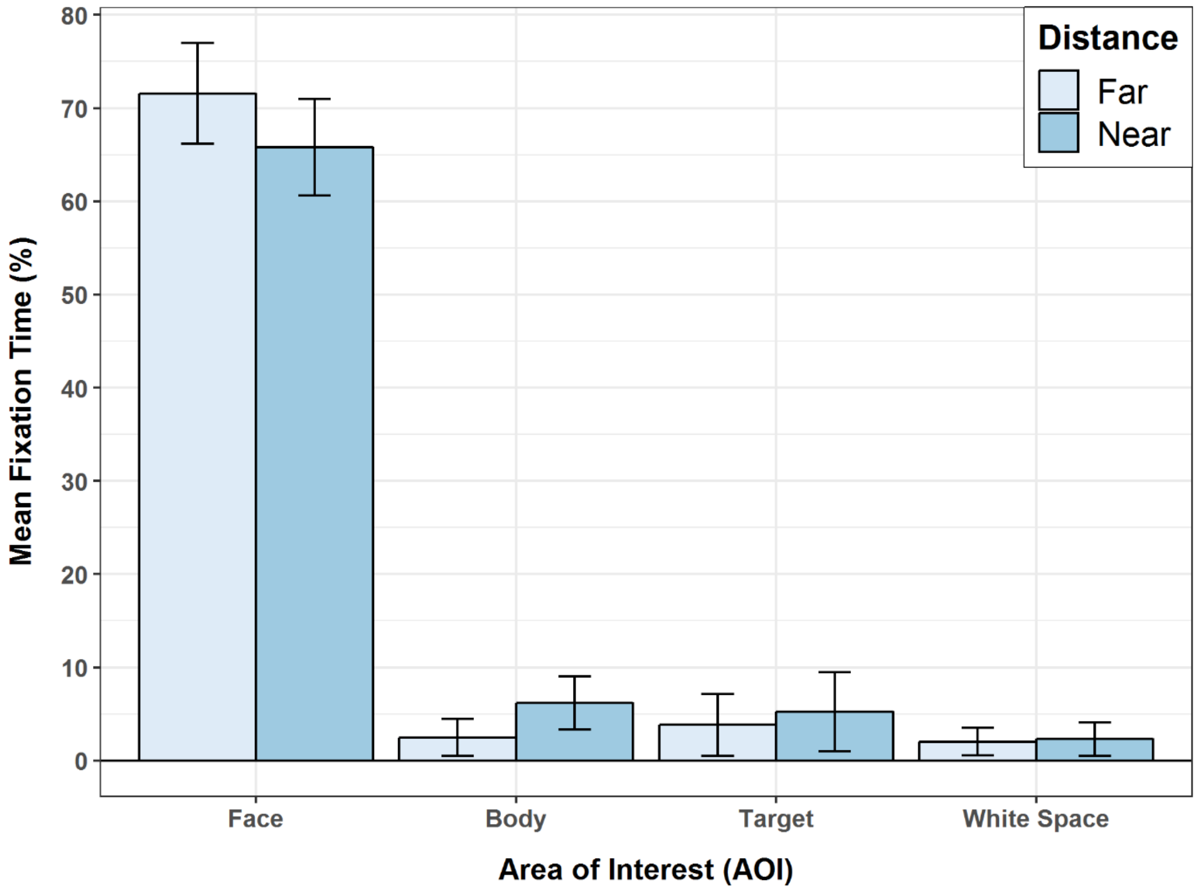

Fig. 7: Mean Fixation Time (%) for Distance (Pointing Far, Pointing Near) and Area of Interest (Face, Body, Target, White Space); Error bars: 95% CI.

The three-way interaction AOI x Distance x Gaze also reached the level of significance, F(3, 114) = 16.48, $\eta^2 = 0.03$, *p[GG]* < .001, (Fig. 8).

To look at the three-way interaction effects in more detail a separate ANOVA was conducted for Pointing Distances Far and Near revealing a significant interaction effect of AOI x Gaze for Near, F(3, 114) = 17.62, η^2= 0.07, *p[GG]* < .001, but not for Far, *p* = .21. Both ANOVAs also revealed a main effect of AOI.

- Pointing Far, main effect of AOI: F(3,114) = 360.64, η^2 = 0.89, *p[GG]* < .001
- Pointing Near, main effect of AOI: F(3, 114) = 234.50, η^2 = 0.83, *p[GG]* < .001

Thus, directional gaze shifts accompanying the pointing bids didn't seem to modulate participants' fixation activity when the Pointing Distance was Far.

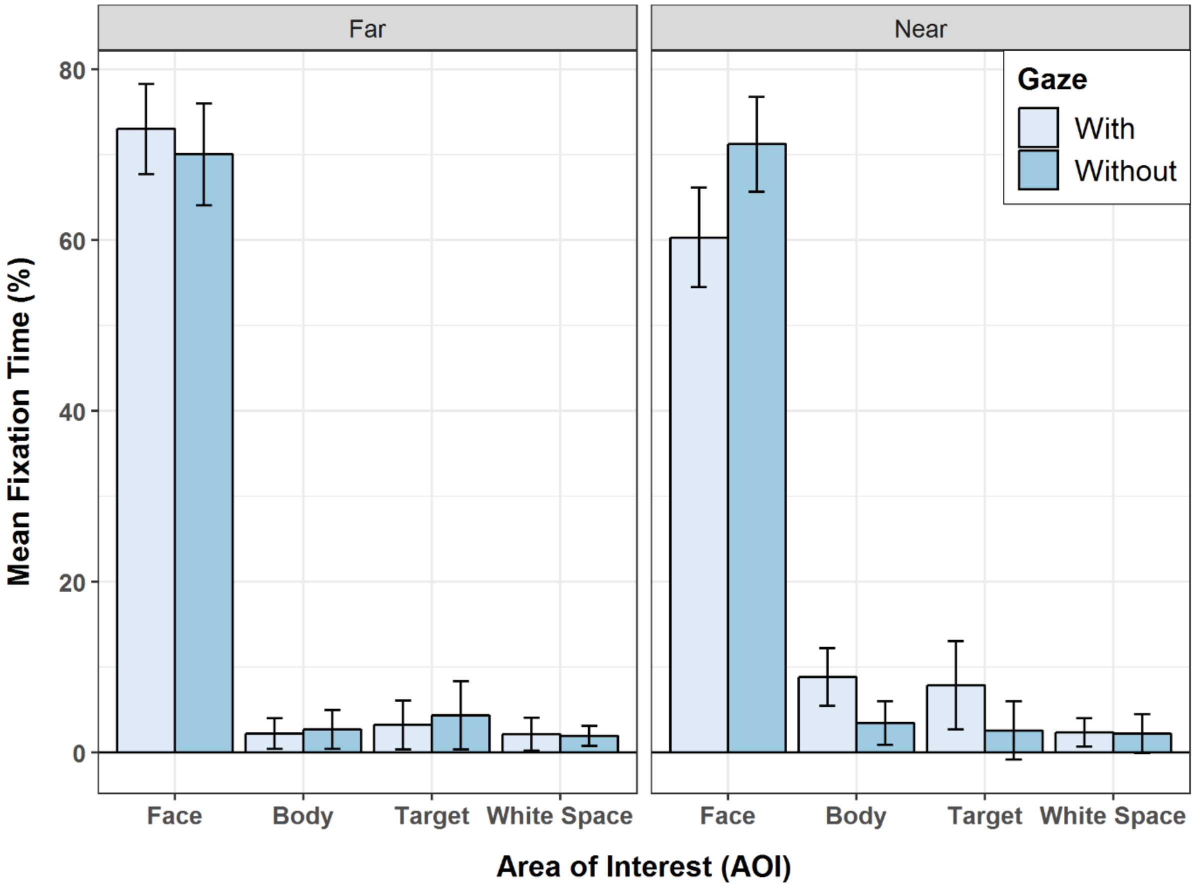

Fig. 8: Mean Fixation Time (%) for Area of Interest (Face, Body, Target, White Space), Distance (Pointing Far, Pointing Near), and Gaze (With Gaze, Without Gaze); Error bars: 95% CI.

By contrast, post hoc paired t-tests (Bonferroni corrected) revealed evidence indicating that Pointing Near accompanied by Gaze seems to be the driving force behind the three-way interaction as under that condition fixational activity, was slightly reduced for the Face, (while still exhibiting the highest amount of fixational activity),

- Near x Face x With Gaze: M = 60.32, 95% CI [54.50, 66.14]
- Near x Face x Without Gaze: M = 71.26, 95% CI [65.70, 76.82]
- Near x Face x With Gaze – Without Gaze: $t(39) = -4.29$, $p < .001$

but increased for the Body,

- Near x Body x With Gaze: M = 8.87, 95% CI [5.48, 12.26]
- Near x Body x Without Gaze: M = 3.50, 95% CI [0.95, 6.05]
- Near x Body x With Gaze – Without Gaze: $t(39) = 4.84$, $p < .001$

and also increased for the Target:

- Near x Target x With Gaze: M = 7.90, 95% CI [2.7, 13.1]
- Near x Target x Without Gaze: M = 2.59, 95% CI [<.001, 6.0]
- Near x Target x With Gaze – Without Gaze: $t(39) = 3.49$, $p = .004$

A similar comparison for the White Space region (With Gaze M = 2.37, 95% CI [0.72, 4.02]; Without Gaze M = 2.23, 95% CI [< .001, 4.50]) didn't reveal any significance, $p > .999$.

Although, it was thus revealed that Pointing Near with Gaze seemed to slightly modulate gaze behaviour of participants in favour of gaze shifts towards the Target, nevertheless in all cases Pointing Distance Far or Near, fixations were always most likely located on the Face as compared to the Body, Target, or White Space.

However, no other effect, particularly the main or interaction effect of Experiment, reached the level of significance, all $p > .16$.

1.3.2 Transition Matrix Full

Based on the sequence of gaze events falling upon one of the 4 AOIs (Face, Body, Target, White Space) the full transition matrix was calculated.

Table 1: Transition Matrix Full.

	Face	Target	White Space	Body	Total
Face	2882	115	67	130	3194
	90.24%	3.61%	2.10%	4.08%	100%
	92.35%C	37.83%C	56.31%C	49.43%C	
Target	76	160	4	25	265
	28.68%	60.38%	1.51%	9.44%	100%
	2.44%C	52.64%C	3.37%C	9.51%C	
White Space	54	4	39	7	104
	51.93%	3.85%	37.50%	6.74%	100%
	1.74%C	1.32%C	32.78%C	2.67%C	
Body	109	25	9	101	244
	44.68%	10.25%	3.69%	41.40%	100%
	3.5%C	8.23%C	7.57%C	38.41%C	
Total	3121	304	119	263	3807

Note: The transition matrix is calculated based on all gaze events in a sequence falling on one of the 4 AOIs Face, Body, Target, and White Space, which denote one row (i) and one column (j) each. Every row and column combination consists of 3 entries. The first entry indicates the transition tallies from AOI i [row] to AOI j [column]. The second entry shows the probability of AOI j [column] following AOI i [row] (%). The third row indicates probability of AOI i [row] preceding AOI j [column] (%C).

Based on Table 1 most transitions in the sequence of gaze events occurred from Face to Face. Further transitions to Face were highly probable:

- Face to Face: 90.24%, Target to Face: 28.68%
- Body to Face: 44.68%, White Space to Face: 51.93%

In contrast transitions to Target, even from Target to Target, were less probable:

- Target to Target: 60.38%, Face to Target: 3.61%
- Body to Target: 10.25%, White Space to Target: 3.85%

Interestingly, Face preceding any of the other AOIs was also highly likely compared to Target preceding any of the other AOIs. Similarly, less likely were Body or White Space AOI preceding any of the other AOIs:

- Face before: Target = 37.83%, Body = 49.43%, White Space = 56.31%
- Target before: Face = 2.44%, Body = 9.51%, White Space = 3.37%
- Body before: Face = 3.5%, Target = 8.23%, White Space = 7.57%
- White Space before: Face = 1.74%, Target = 1.32%, Body = 2.67%

Further, Face preceding Face was much higher than Target preceding Target, Body preceding Body or White Space preceding White Space:

- Face before Face: 92.35%, Target before Target: 52.64%
- Body before Body: 38.41%, White Space before White Space: 32.78%

1.3.3 Transition Matrix Collapsed

Based on the collapsed sequence of gaze events by disregarding transitions within AOIs, but only counting transitions between AOIs the collapsed transition matrix was calculated. This is the reason why the collapsed transition matrix always shows 0 when row and column are the same AOI.

Table 2: Transition Matrix Collapsed.

	Face	Target	White Space	Body	Total
Face	0	115	67	130	312
	0.00%	36.86%	21.48%	41.67%	100%
	0.0%C	79.87%C	83.75%C	80.25%C	
Target	76	0	4	25	105
	72.39%	0.00%	3.81%	23.81%	100%
	31.8%C	0.0%C	5.0%C	15.44%C	
White Space	54	4	0	7	65
	83.08%	6.16%	0.00%	10.77%	100%
	22.6%C	2.78%C	0.0%C	4.33%C	
Body	109	25	9	0	143
	76.23%	17.49%	6.30%	0.00%	100%
	45.61%C	17.37%C	11.25%C	0.0%C	
Total	239	144	80	162	625

Note: The collapsed transition matrix is calculated solely based on transitions between but not within AOIs Face, Body, Target, and White Space, which denote one row (i) and one column (j) each. The first entry of every cell shows the number of transitions from AOI i [row] to AOI j [column]; the second entry displays probability of AOI j [column] following AOI i [row] (%); and the third entry indicates probability of AOI i [row] preceding AOI j [column] (%C).

Based on Table 2 most transitions between AOIs occurred from Face to Body (130). Further, following Target, Body and White Space always Face was most likely to occur next rather than Target or one of the other AOIs:

- Target to: Face = 72.39%, Body = 23.81%, White Space = 3.81%
- Body to: Face = 76.23%, Target = 17.49%, White Space = 6.30%
- White Space to: Face = 83.08%, Target = 6.16%, Body = 10.77%

Interestingly, Face preceding any of the other AOIs was also much more likely compared to Target, Body or White Space AOIs preceding any of the other AOIs.

- Face before: Target = 79.87%, Body = 80.25%, White Space = 83.75%
- Target before: Face = 31.8%, Body = 15.44%, White Space = 5.0%
- Body before: Face = 45.61%, Target = 17.37%, White Space = 11.25 %
- White Space before: Face = 22.6%, Target = 2.78%, Body = 4.33%

In summary, the transition matrix tallies for full and collapsed transition matrix as well as the probability calculations (see paragraph 1.3.4) indicate that fixations fell most likely on the face of the experimenter, rather than on the target location of displayed directional stimuli.

1.3.4 Probability of Gaze Movements to Areas of Interest

Results from the calculation of fixation probability of each of the four AOIs for the first four fixations and for each of the four types of social cue (all Distance and Gaze combinations) yielded similar results (Fig. 9).

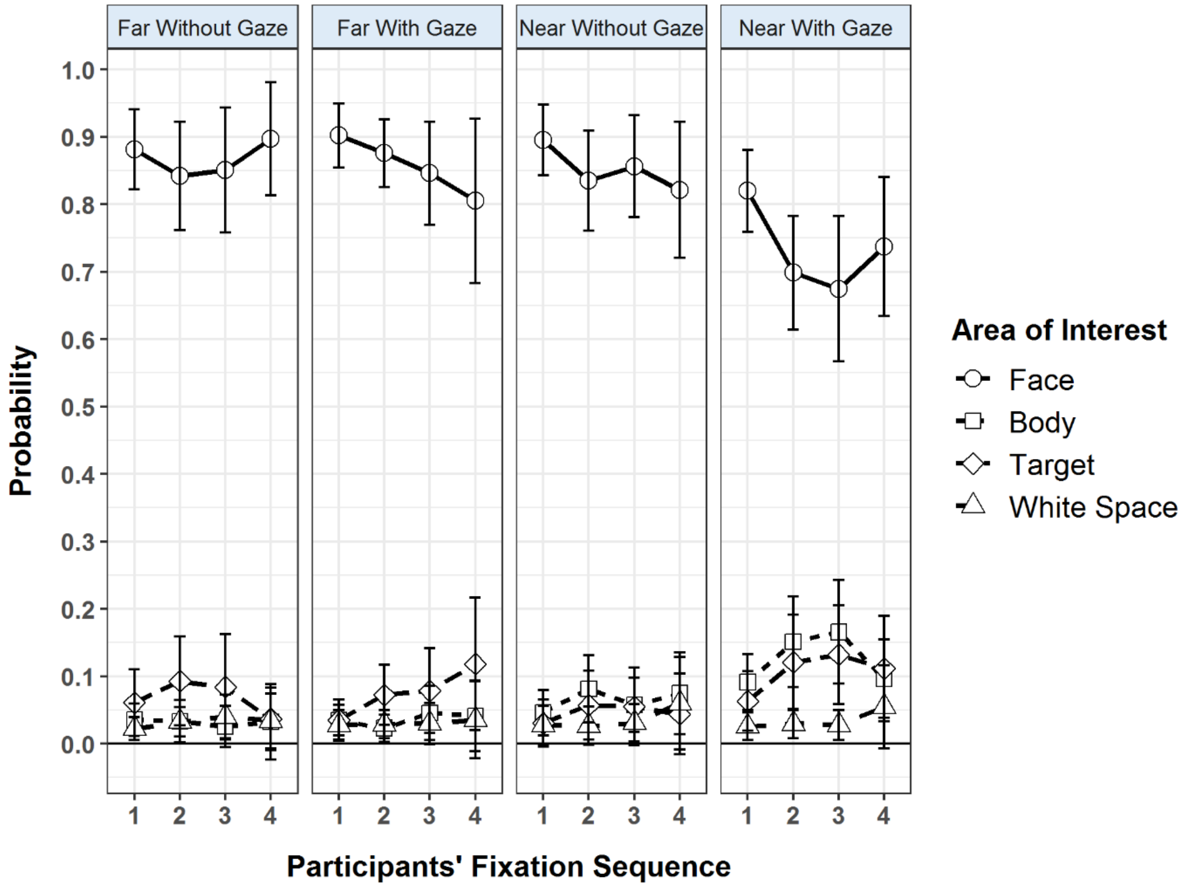

Fig. 9: Probability for the first 4 fixations falling on each of the Areas of Interest Face, Body, Target, or White Space based on each Type of directional stimulus (Far, Far + Gaze, Near, Near + Gaze); Error bars: 95% CI.

Figure 9 shows that independent of the type of directional stimulus the first four fixations seemed to fall substantially more likely (no overlap of the 95% CI, calculation based on the correction method from Morey (2008)) on the Face rather than the Target or any of the other AOIs. Furthermore, Target, Body and White Space AOI appear similarly less probable. Thus, it is unlikely that participants first briefly directed their gaze towards the distractor Target and then rapidly disengaged and fixated back onto the Face for a longer time period, which couldn't have been ruled out on the above fixation time analysis alone.

1.4 Discussion

This study aimed at examining gaze response towards irrelevant directional cues during an interactive real-life scenario. For that purpose, an experimenter conversed with participants in dyadic pairs while performing pointing gestures and gaze shifts at scripted times. Reaction of participants to these signals was measured via mobile eye-tracking. It has been suggested by previous studies that social directional stimuli may lead to attentional capture, i.e. reflexive gaze following (Friesen & Kingstone, 1998; Friesen, Moore, & Kingstone, 2005; Gluckman & Johnson, 2013). However, it has yet to be determined whether in real interactive situations, i.e. when the observed stimuli are based on movements of an experimenter instead of pictures, this reflexive orienting of attention would persist. On the other hand, it has been proposed that attentional orienting is influenced by top-down cognitive factors, whereas context (Wiese, Zwickel, & Müller, 2013) as well as prior beliefs about the interaction partner's mental states influences reflexive shifting of attention (Teufel et al., 2010; Wiese et al., 2012). Thus, the goal of this study was to investigate whether participants would reflexively follow pointing gestures and gaze cues of an experimenter that are irrelevant to the topic of conversation.

The present results revealed that during an ongoing conversation participants' gaze fell most likely on the face of the experimenter in spite of the directional cues that were presented. This occurred irrespective of the strength or type of directional cue, i.e. gaze shift + pointing or only pointing, or the culture of all interaction partner, i.e. all either from Japan or from Germany (Fig. 5-8). Importantly, the full and collapsed transition matrix analysis provided similar evidence as transitions from and to the AOI "Face" were much more likely than transitions to the AOI "Target" - indexing the location of directional cues - or to the AOI "Body" as well as "White Space" - indicating the experimenters' body and the rest of the environment respectively (Table 1, Table 2). Another AOI sequence analysis of the first 4 fixations after onset of directional cue provided insight on whether participants' fixations possibly occurred first towards the observed directional cues and only then towards the face of the experimenter as has been proposed by Theeuwes and colleagues' "capture and rapid disengagement" hypothesis (e.g., Theeuwes, 2010; Theeuwes, Atchley, & Kramer, 2000).

Similarly, to the transition matrix and fixation time analysis however this fine-grained sequence analysis also showed evidence that the first fixations were substantially more likely to fall on the task-relevant face of the experimenter (Fig. 9.) rather than the task-irrelevant location of directional cues. Interestingly, taking the direction of the pointing bids into account, either pointing near or far away, didn't make a difference in the results as fixations remained largely on the face (Fig. 7, Fig. 8) independent of the pointing distance.

Nevertheless, when experimenters' pointing near bids were accompanied by gaze shifts in the same direction gaze reaction of participants was slightly altered as fixations seemed to occur slightly more often towards the body or target (Fig. 8). As this came to pass solely for the pointing near + gaze condition, it is reasonable to assume that this signal was slightly more potent than any of the other combination of signals in terms of capturing attention. Though, one must keep in mind that the face was still the main fixation spot.

This might mean that in the context of a natural dyadic conversation directional Gaze + Pointing Near bids appear more centred, and thus more shared, as compared to, for example, Gaze + Pointing Far bids that might indicate a wider space or just Pointing Near bids that aren't accompanied by gaze and therefore might lack strength. In this vein, attentional orienting of participants towards distracting directional cues might depend not only on the strength of the directional stimuli, accompanied or not accompanied by gaze shifts, but also on the perceived "social sharedness" of the cue, i.e. as soon as an adequate joint context is created.

All the evidence thus provided may give rise to the notion that in real life scenarios no purely reflexive attentional capture occurs towards social directional cues as has been previously demonstrated with 2D picture stimuli. Nevertheless, the underlying reason for the lack of attentional capture may be due to the nature of the provided stimuli. These stimuli were embodied pointing bids sometimes accompanied by gaze shifts of a present experimenter that had little bearing towards the content of the conversation and may thus have been regarded as meaningless distractors. However, previous studies have proposed that even distractor signals elicit attentional capture and that indeed the reflexive orienting of attention can't be successfully suppressed when a social signal, such as for example, a gaze shift, is perceived. (Ristic & Kingstone, 2005).

This has been also explored in non-social studies on spatial attention tasks, such as the additional-singleton paradigm in visual search, where findings have demonstrated that meaningless, task-irrelevant but salient distracter signals may reflexively capture attention, and even summon eye movements towards the distractor location (Theeuwes, 1994, 2004). Attentional capture can occur even through non-consciously perceived subliminal cues (Mulckhuyse & Theeuwes, 2010), for review). Most importantly, previous evidence has demonstrated that following a perceived gaze cue an involuntary oculomotor, gaze-following response from the observer's side emerges, which suggests the occurrence of a certain degree of oculomotor mimicry during interaction (Kuhn & Kingstone, 2009; Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002).

On the other hand, there is also evidence that attentional capture in visual search can be modulated by top-down factors (e.g., featural or dimensional set: Müller et al., 2009; Töllner, Müller, &

Zehetleitner, 2011; Wykowska & Schubö, 2010; Wykowska & Schubö, 2011). Also, in the social attention literature, gaze following has been reported to be reduced, for example, when observers believed that an avatar was blind towards the environment, for example because it was perceived to be wearing non-transparent goggles (Teufel et al., 2010). Also, beliefs about the intentionality of agents have been shown to influence social attentional orienting (Wiese et al., 2012; Wykowska et al., 2014). Finally, also social context has been shown to modulate gaze cueing effects even in more naturalistic scenarios (Kompatsiari et al., 2018).

Similarly, this may explain the findings of this study as there is evidence which supports the hypothesis that top-down factors play a substantial role in modulating attentional capture.

It is important to emphasize that in this study stimuli were embodied by a human experimenter during a natural conversation in contrast to similar previous experiments which made use of pictures and a limited interaction context. This is of note because it has been argued that only embodied stimuli as well as the give-and take of a real interaction elicits certain social cognitive mechanisms. In that context alone participants put themselves in the position to really engage with an interaction partner and consequently only then make use of the necessary cognitive mechanisms, which may be a reason for differences between ecologically valid interaction studies and traditionally more static studies (Hayward, Voorhies, Morris, Capozzi, & Ristic, 2017; Schilbach et al., 2013; Skarratt, Cole, & Kingstone, 2010; Skarratt et al., 2012) Accordingly, natural interaction and more ecologically valid experimental protocols might cast a different light on social cognitive mechanisms than screen-based “observational” protocols.

However, an alternative explanation might account for the pattern of results in this study: perhaps instead of top-down modulation of reflexive orienting of attention, certain cognitive mechanism such as attentional capture might not have been activated at all.

Specifically, the activation of automatic, bottom-up attentional orienting might only arise when a sufficient level of stimulus strength/energy, specific context and previous experiences suffice to reach a pre-determined threshold of neural activity(Cole, Smith, & Atkinson, 2015; Zehetleitner, Hegenloh, & Müller, 2011; Zehetleitner, Krummenacher, & Müller, 2009).
In the context of computer experiments with static stimuli, this activation level may be easily reached with directional cues, as the stimulus is salient enough due to being isolated. In contrast, during natural interaction with a real human pointing bids and gaze shifts alone might be insufficient to trigger attentional capture in the face of the large amalgamation of signals that are implicitly elicited by a human being. These signals compete against each other and thus make the single or double directional signal(s) less salient.

1.4.1 Summary and Conclusion

As a conclusion, during interaction with an embodied interaction partner in a naturalistic set-up attentional orienting as measured by oculomotor response might not arise based on irrelevant directional cues. On the contrary, participants seem to remain focused on the face of the interaction partner to listen to the conversation instead of being distracted by irrelevant pointing gestures or gaze shifts. Interestingly, this effect seems to be culturally robust as there was no difference between German and Japanese participants.

Chapter 2 Intercultural Gaze Patterns

2.1 Introduction

Research Question:

Do participants show different gaze patterns while listening, being asked a question and answering and does this behaviour vary based on culture?

Result:

Japanese and German participants show similar gaze patterns while listening and being asked a question, that is they focus mostly on the face. The same seems to be true also for answering German participants. However, Japanese participants seem to avoid looking at the face equally as much as they look at the face of their interaction partner when answering.

Based on evidence collected in the previous chapter (Chapter 1) it was shown that gaze during interaction remains mostly focused on the partner's face, which seems to be the case independent of culture. However, so far, we have investigated gaze behaviour only while listening. Furthermore, it has been reported that gaze patterns change from listening to being asked a question to answering. Specifically, people seem to look more at interaction partner when listening rather than when speaking (Cook, 1977; Kendon, 1967) and they seem to look away when asked a question, particularly when the question is difficult (see review Cook, 1977). Importantly, it has been suggested that culture may play an important role in modulating gaze behaviour while answering. In that context Japanese participants, for example, seem to avert gaze more than Canadian or Trinidadian participants while answering (McCarthy, Lee, Itakura, & Muir, 2006).

Typical for many of the above-mentioned interaction studies (Chapter 1, Kendon, 1967; McCarthy et al., 2006), experimenters were tasked to interact with participants, often in an interview like situation. Hence, it was possible to simulate a real-world interaction in a naturalistic setting. In line with Schilbach and colleagues' (2013) second-person neuroscience approach, this is an advantage above similar studies investigating social behaviour in laboratory settings where participants merely react to 2D stimuli shown on a screen. Being embedded in a real social situation, as opposed to observing social stimuli on a screen, seems to trigger marked differences not only in behaviour, but also in the underlying cognitive resources that are recruited to process stimuli (Cole et al., 2016; Schilbach et al., 2013). In this vein and similar to the experiment of Chapter 1, the present study sought to replicate a naturalistic real-world scenario with all the advantages implied therein by immersing participants in a real dyadic interaction with an experimenter.

2.1.1 Aim of Study

The goal of the present study, which was based on an experiment of Kajopoulos (2015), was to explore differences and similarities between gaze behaviour in German and Japanese cultures during various types of interaction. Specifically, I aimed at understanding whether the face of an interaction partner is the default “home” position in both cultures, Germany and Japan, and I intended to test whether this is modulated by the type of interaction, i.e., listening, being asked a question, answering. To that purpose, I used a mobile eye-tracking system to analyse gaze patterns during a conversation with an experimenter who followed a script.

2.2 Materials and Methods

As this was the second part of the experiment described in Chapter 1, information about participants, apparatus and setup, as well as calibration procedure and cover story are the same as in Chapter 1 – Materials and Methods.

2.2.1 Procedure – Interaction Script

For the conversation with the participants a scripted protocol was used from the experimenter's side.

After talking about the experiment (L1=Listening trial 1), the experimenter asked a question (Q1). Following a participant's answer, which was spontaneous and natural from the participants' side (A1), the experimenter talked again (L2) and so the trials repeated until A5. [L1, Q1, A1, L2, Q2, A2, L3, Q3, A3, L4, Q4, A4, L5, Q5, A5...].

As during the listening trials [L1 – L5] experimenters talked for a longer time (approximately 30 seconds each) than during the questioning trials (approximately 3 seconds each) additional questioning trials were needed to record sufficient amount of gaze events, thus the interaction script included more question and answer phases relevant for the experiment [A6-A25; Q6-Q25]. If a participant in anticipation of the question already answered a question in a previous trial the experimenter skipped asking the question.

2.2.2 Data Analysis

Gaze data was pre-processed with the BeGaze software (SensoMotoric Instruments GmbH, Germany) to tag fixation events with their corresponding Area of Interest (Face, Body, White Space) label (Fig. 10).

Fixation Time (sum of durations of all fixations) per participant, for the different Interaction Types: listening (phases: L1-L5), being questioned (phases: Q1-Q25) and answering (phases: A1-A25), for the Areas of Interest (AOI): Face, Body and White Space (= rest of environment, everything that is not Face or Body) were extracted from BeGaze. Data analysis on mean fixation time [%], i.e. mean (fixation time / trial duration) was performed.

Mixed-design Analyses of Variance (ANOVAs) were conducted for all measures with Interaction Type and AOI as within-subjects factor and Culture as between-subjects- factor. Mauchly's test for sphericity was performed and if significant, the Greenhouse-Geisser corrected p-values *p[GG]* were calculated. 95% confidence intervals were calculated based on the correction method from Morey (2008), which takes within-subjects factors into account.

Fig. 10: Experimental Procedure, Area of Interest (AOI) Face (blue), Body (grey), everything that is not AOI Face or Body is Area of Interest White Space.

2.3 Results

The mixed-design ANOVA revealed a significant main effect of Area of Interest (Face, Body, White Space), $F(2, 78) = 308.30$, $\eta^2 = 0.83$, $p[GG] < .001$ (Fig. 11) and a significant main effect of Interaction Type (Listening, Questioned, Answering), $F(2,78) = 63.02$, $\eta^2 = 0.05$, $p[GG] < .001$ (Fig. 12). Post-hoc comparisons using paired t-tests (Bonferroni-corrected for multiple comparisons) revealed that significantly more fixation activity was centred on the Face AOI (M = 50.88, 95% CI [46.74, 55.02]) than any of the other two AOIs, the Body (M = 6.78, 95% CI [3.43, 10.13]), $t(40) = 15.46$, $p < .001$, and the White Space (M = 10.31, 95% CI [7.67, 12.95]), $t(40) = 16.51$, $p < .001$, with no difference between the latter two AOIs, $p > .15$ (Fig. 11).

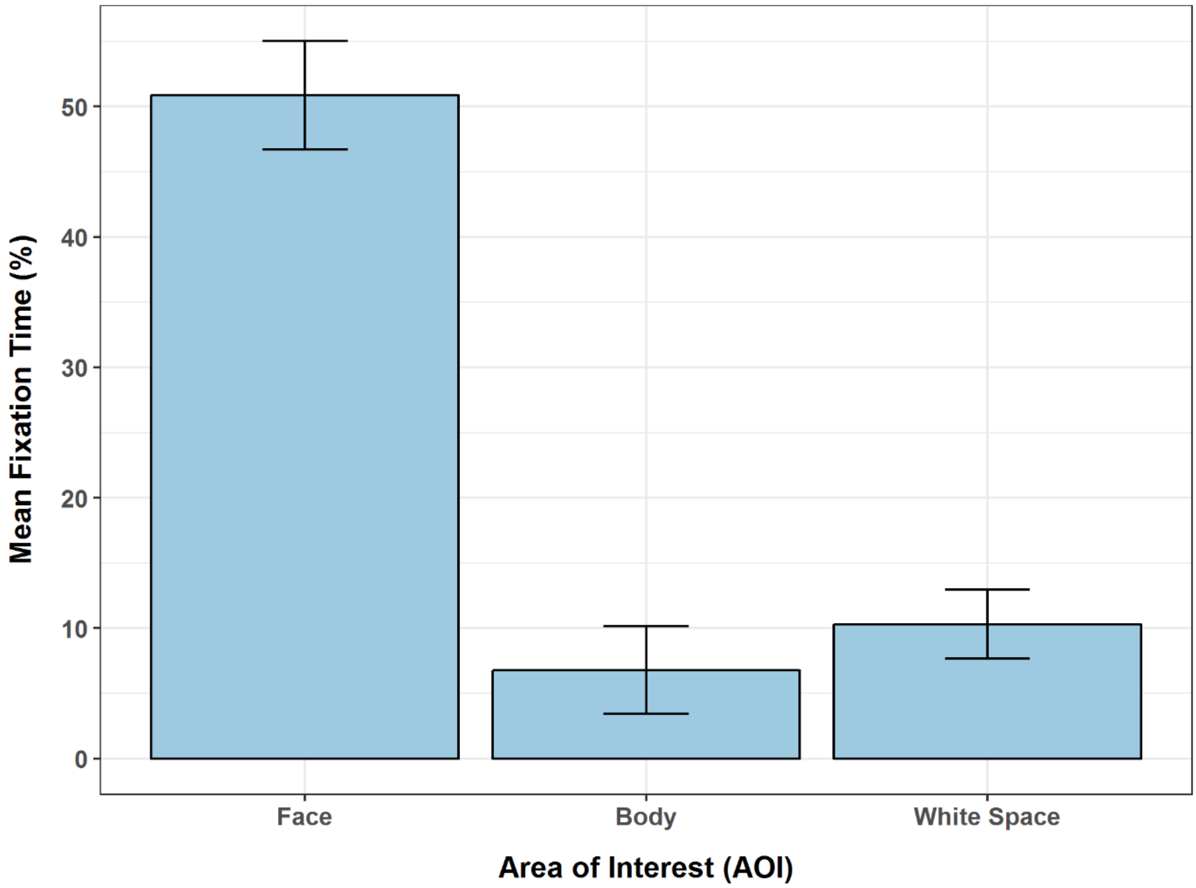

Fig. 11: Mean Fixation Time (%) for Area of Interest (Face, Body, White Space); Error bars: 95% confidence interval (CI).

Furthermore, post-hoc tests showed significantly more fixation activity occurred while being questioned (M = 25.20, 95% CI [24.53, 25.87]) than while listening (M = 24.13, 95% CI [23.54, 24.72]), $t(40) = 3.26$, $p = .007$, or answering (M = 19.80, 95% CI [18.88, 20.72]), $t(40) = 8.84$, $p < .001$. Also, there was significantly more fixational activity while listening than while answering, $t(40) = 7.81$, $p < .001$ (Fig. 12).

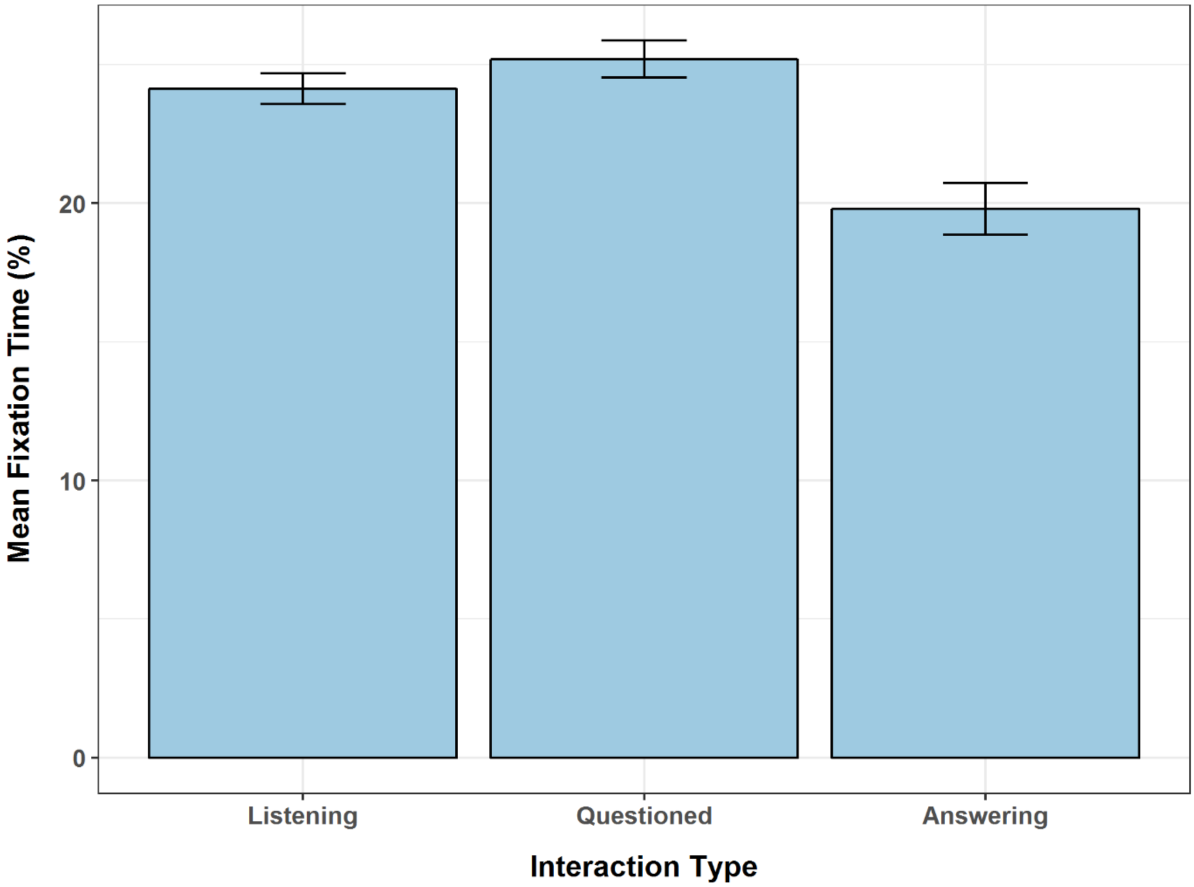

Fig. 12: Mean Fixation Time (%) for Interaction Type (Listening, Questioned, Answering); Error bars: 95% CI.

Furthermore, results revealed a significant interaction effect between AOI and Interaction Type, *F*(4, 156) = 109.29, η^2 = 0.44, *p[GG]* < .001 (Fig. 13).

Concerning the AOI x Interaction Type effect, post-hoc comparisons (paired t-tests, Bonferroni-corrected for multiple comparisons) showed that significantly less fixation activity was centred on the Face AOI during the Answering phase (M = 33.78, 95% CI [28.77, 38.79]) than during both the Listening (M = 66.58, 95% CI [62.65, 70.51]), *t*(40) = -13.42, *p* < .001, and the being Questioned phases (M = 64.62, 95% CI [60.50, 68.74]), *t*(40) = -12.40, *p* < .001. Similar amounts of fixation activity were centred on the Face, during the Listening and being Questioned phases, *p* > .999. On the other hand, participants seemed to fixate the Body AOI with similar amounts while listening (M = 3.63, 95% CI [0.81, 6.45]), being questioned (M = 5.26, 95% CI [2.27, 8.25]) and answering (M = 9.0, 95% CI [5.25, 12.75]), all *p* > .09. In contrast, for the White Space AOI fixational activity differed significantly between all 3 phases, with more fixation activity occurring during answering (M = 16.61, 95% CI [13.28, 19.94]) than while being questioned (M = 5.72, 95% CI [3.60, 7.84]), $t(40) = 8.11$, $p < .001$, and listening (M = 2.18, 95% CI [0.57, 3.79]), $t(40) = 10.0$, $p < .001$, and also more fixation activity occurring while being questioned than while listening, $t(40) = 6.17$, $p < .001$. Most interestingly, the typical pattern of higher fixation activity on the Face than on both the Body (Listening: $t(40) = 20.75$, Questioned: $t(40) = 19.03$, Answering: $t(40) = 6.85$) and White Space (Listening: $t(40) = 27.02$, Questioned: $t(40) = 21.89$, Answering: $t(40) = 5.09$), all $p < .001$, and no difference in fixation amount for the latter two AOIs occurred for all three interaction types, all p >. 12, (Fig. 13).

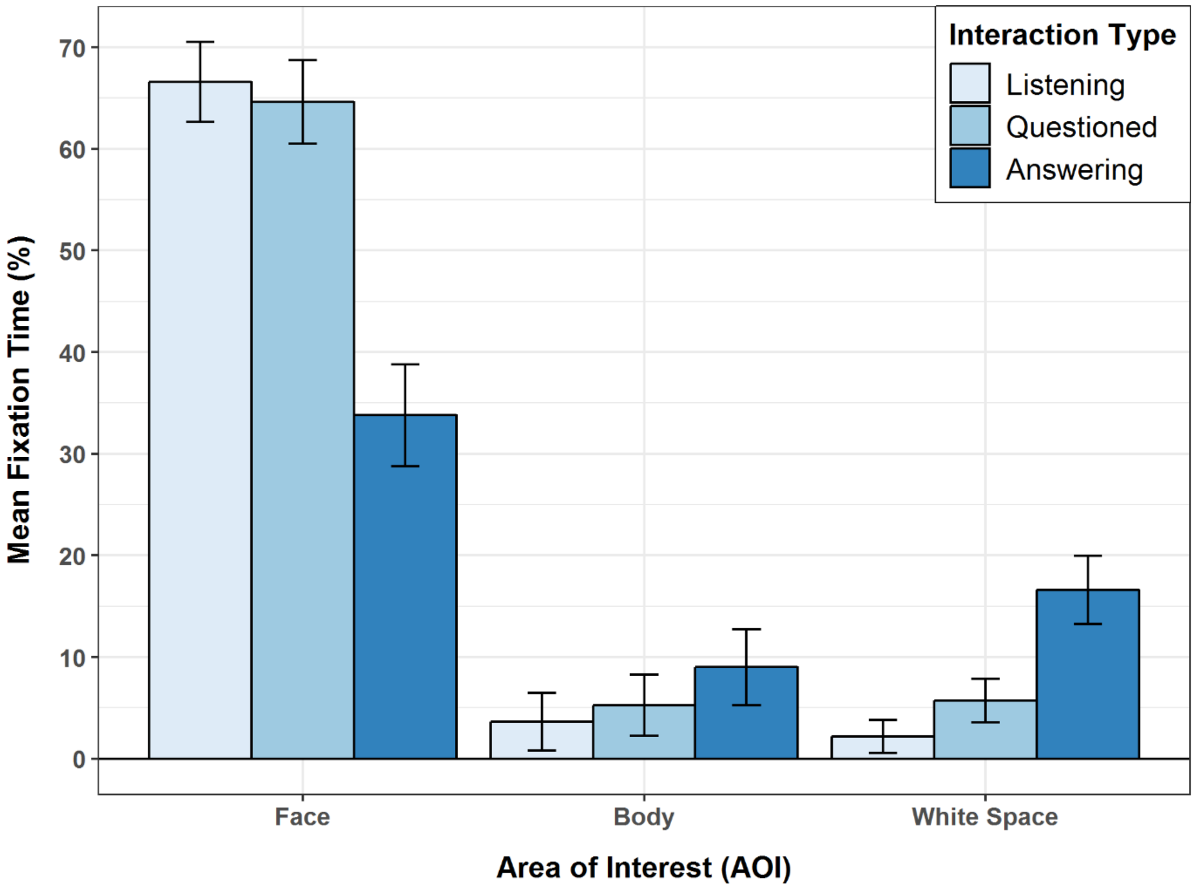

Fig. 13: Mean Fixation Time (%) for Area of Interest (Face, Body, White Space) and Interaction Type (Listening, Questioned, Answering); Error bars: 95% CI.

Finally, the three-way interaction among AOI, Interaction Type and Country was significant, $F(4,156) = 4.16$, $\eta^2 = 0.03$, $p[GG] = .01$ (Fig. 14).

Neither, the main effect of Country nor the interaction effects of Country x Interaction Type or Country x AOI reached the level of significance, all $p > .24$.

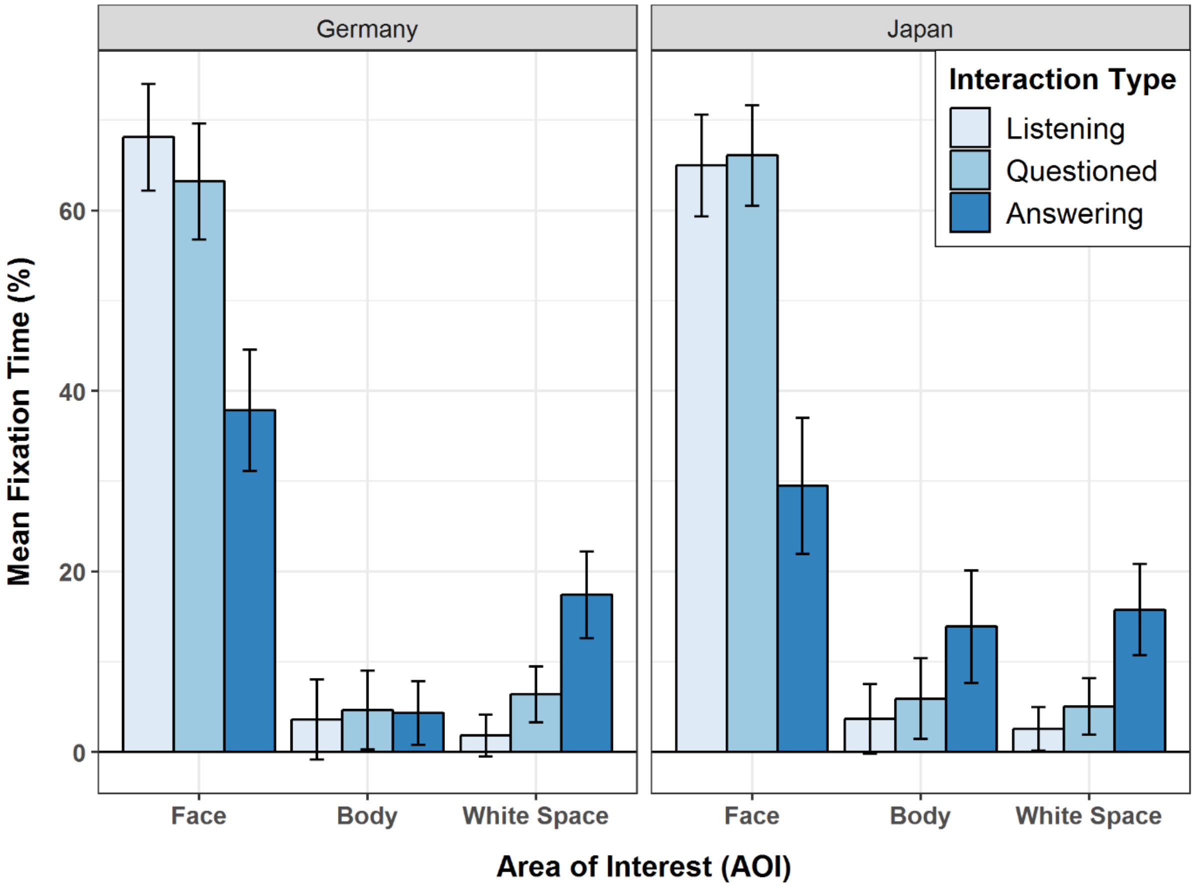

Fig. 14: Mean Fixation Time (%) for Area of Interest (Face, Body, White Space), Interaction Type (Listening, Questioned, Answering) and Country (Germany, Japan); Error bars: 95% CI.

To further analyse this three-way interaction a separate ANOVA for each Interaction Type was performed.

As expected from results of above t-tests all separate ANOVAs for “Listening”, “Questioned” and “Answering” yielded a significant main effect of AOI (Listening: $F(2,78) = 482.17$, $\eta^2 = 0.92$, $p[GG] < .001$; Questioned: $F(2,78) = 358.49$, $\eta^2 = 0.89$, $p[GG] < .001$; Answering: $F(2,78) = 32.58$, $\eta^2 = 0.42$, $p < .001$) (see Table 3).

Table 3: Mean Fixation Time (%) and 95% CI for Country (Germany, Japan), Area of Interest (Face, Body, White Space), Interaction Type (Listening, Questioned, Answering).

Country	AOI	Interaction Type	Mean Fixation Time (%)	95% CI
Germany	Face	Listening	68.11	[62.22, 74.0]
Germany	Face	Questioned	63.22	[56.77, 69.67]
Germany	Face	Answering	37.86	[31.16, 44.56]
Germany	Body	Listening	3.61	[< .001, 8.07]
Germany	Body	Questioned	4.64	[0.27, 9.01]
Germany	Body	Answering	4.34	[0.80, 7.88]
Germany	White Space	Listening	1.82	[< .001, 4.15]
Germany	White Space	Questioned	6.38	[3.28, 9.48]
Germany	White Space	Answering	17.41	[12.63, 22.19]
Japan	Face	Listening	64.97	[59.34, 70.60]
Japan	Face	Questioned	66.10	[60.52, 71.68]
Japan	Face	Answering	29.50	[21.96, 37.04]
Japan	Body	Listening	3.65	[< .001, 7.50]
Japan	Body	Questioned	5.91	[1.44, 10.38]
Japan	Body	Answering	13.89	[7.65, 20.13]
Japan	White Space	Listening	2.56	[0.14, 4.98]
Japan	White Space	Questioned	5.02	[1.89, 8.15]
Japan	White Space	Answering	15.77	[10.71, 20.83]

Interestingly, solely the ANOVA for the Interaction Type “Answering”, revealed the AOI x Country interaction to be significant, $F(2,78) = 4.20$, $\eta^2 = 0.08$, $p = .02$. Based on post-hoc tests (independent, Welch-Two-Sample-t-tests, Bonferroni-corrected for multiple comparisons), significantly more fixation activity was centred on the Body for Japanese participants (M = 13.89, 95% CI [7.65, 20.13]) than for German participants (M = 4.34, 95% CI [0.80, 7.88]), $t(27.51) = 2.87$, $p = .02$. The smaller amount of fixational activity centred on the Face for Japanese participants (M = 29.50, 95% CI [21.96, 37.04]) as compared to German participants (M = 37.86, 95% CI [31.16, 44.56]) didn’t reach the level of significance, $p = .30$. Also, a similar amount of fixation activity was centred on White Space for both, German (M = 17.41, 95% CI [12.63, 22.19]) and Japanese participants (M =15.77, 95% CI [10.71, 20.83]), $p > .999$.

To examine whether, particularly during the Answering Interaction Type, participants typically looked above all at the face of the interaction partner, we compared mean fixation time of Face to mean fixation time of conjoint Body/White Space and conducted an ANOVA for the Interaction Type Answering accordingly, which revealed the main effect of AOI (Face M = 33.78, 95% CI [28.13, 39.43]; Body/WS M = 25.60, 95% CI [19.95, 31.25]), $F(1,39) = 4.43$, $\eta^2 = 0.08$, $p = .04$, and the AOI x Country interaction to be significant, $F(1,39) = 4.40$, $\eta^2 = 0.08$, $p = .04$ (Fig. 15).

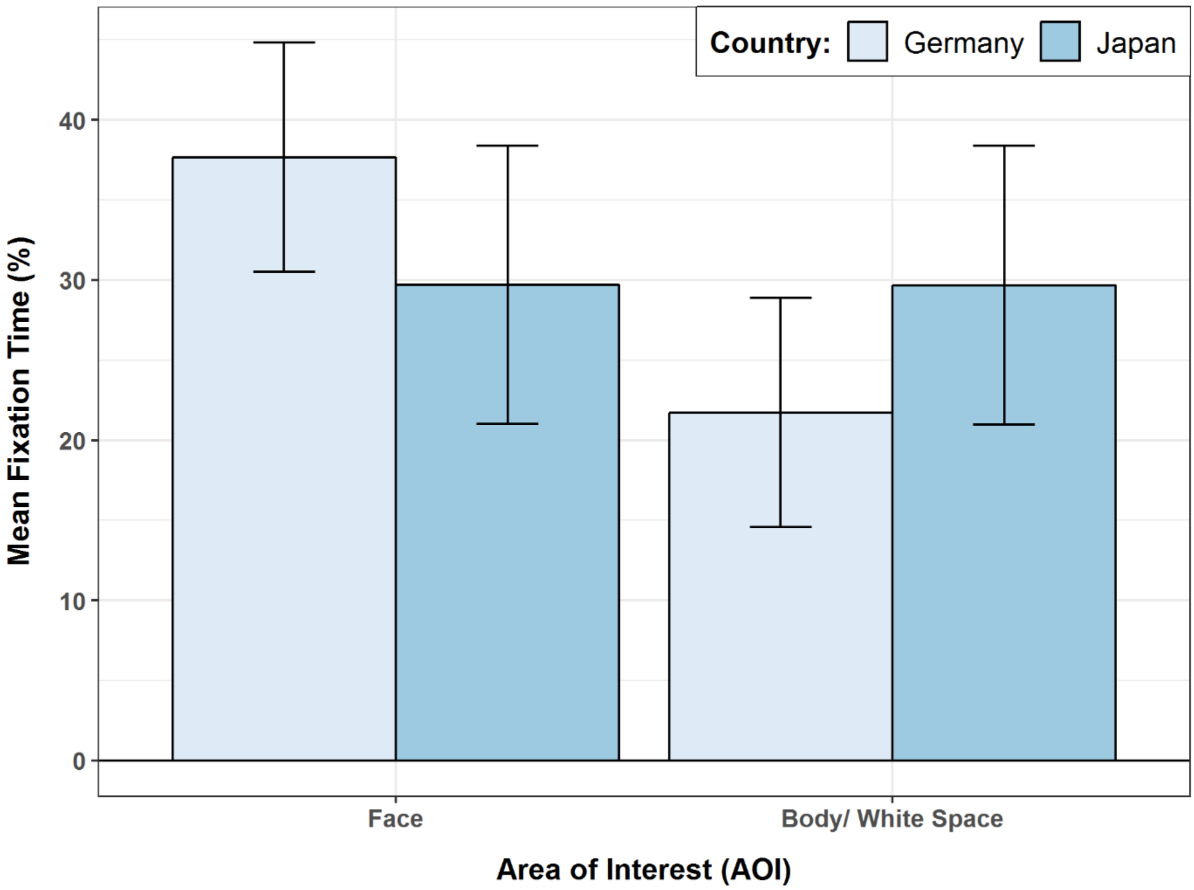

Fig. 15: Mean Fixation Time (%) for Area of Interest (Face, Body/White Space), Interaction Type – Answering, and Country (Germany, Japan); Error bars: 95% CI.

Post-hoc t-tests (paired t-test and for comparison between countries independent Welch-Two-Sample-t-test, both Bonferroni-corrected for multiple comparisons) revealed that Japanese participants fixated similarly on the Face (M = 29.70, 95% CI [21.0, 38.40]) and the rest of the environment (Body/White Space: M = 29.68, 95% CI [20.98, 38.38]), $p > .999$, whereas German participants looked more at the Face (M = 37.67, 95% CI [30.51, 44.83]) than at other parts of the environment (Body/White Space: M = 21.71, 95% CI [14.55, 28.87]), $t(20) = 3.29$, $p = .01$. However, no difference in fixation activity between German and Japanese participants was found for Face nor for Body/White Space, all $p > .18$.

2.4 Discussion

The aim of the present study was to examine gaze patterns of participants during different types of interaction (listening, being asked a question, answering) for two different cultures (Germany, Japan). I hypothesized that there might be a difference in gaze behavior based on the type of interaction, dependent also on the specific culture. For the purpose of my study, I conducted a real-life interaction scenario between participants and a trained experimenter. The experimenter followed a previously learned script to converse with participants.

Results showed that in general participants looked more towards the face than towards the body of the experimenter or the rest of the environment (Fig. 11), which is also similar to the evidence obtained in Chapter 1. Importantly, this pattern occurred for any of the three interaction types (Listening, being Questioned, Answering) (Fig. 13). Thus, it seems clear that participants engaged with the experimenter. This is in line with previous work on gaze behavior during social encounters which suggests that the face is the most relevant feature during a conversation and that people strive to present this "ordinary" gaze behavior while engaged if there is no other competing activity going on (Rossano, 2013). However, there were also differences of gaze behavior based on the three interaction types (Fig. 13). Results indicated that participants looked more at the face of the interaction partner while listening or while being questioned, than while answering (Face: L>A, Q>A) and showed an inverse pattern for the rest of the environment (White Space: L<A, Q<A). On the other hand, participants seemed to look in equal amounts at the face of the experimenter while listening and being questioned (Face: L = Q), but looked more at the rest of the environment while being questioned than while listening (White Space: L< Q). Also, although there may be little difference in regards to fixations on the body of the experimenter for any of the interaction types, there still seems to exist an overall trend which is similar to the gaze pattern for the rest of the environment: participants tended to look more at the body while answering a question than while listening or being questioned (Fig. 13). Previous studies have shown that a thinking process during interaction elicits gaze aversion, i.e. gaze movements away from the interaction partner's face (Doherty-Sneddon & Phelps, 2005; Kinsbourne, 1972; McCarthy et al., 2008). Hence, it is reasonable to assume that while being questioned, and especially while answering questions, participants may need to think about their answer or even recall information from memory to be able to answer. Consequently, they may need to look away from their interaction partner's face.

Interestingly, the larger number of gaze fixations on the face of an interaction partner as opposed to other places in the environment seemed to be modulated not only by interaction type, but also by culture. Although while listening and being questioned both German and Japanese participants showed similar gaze patterns of looking mostly towards the face of an interaction partner, during the answering phase a different gaze pattern seemed to arise, particularly for Japanese participants (Fig. 14, Fig. 15). Including the results for merged fixations from Body and White Space AOIs, it becomes clear that where German participants consistently continued to gaze more at the face of the interaction partner than at the rest of the environment during the answering phase, Japanese participants show a different pattern by looking with equal amounts at the face and at the body and the rest of the environment combined (Fig. 15). In fact, while speaking Japanese participants seemed too have shifted their gaze from the face to the body of the interaction partner, as indicated by the results that showed increased fixation activity on the Body AOI for Japanese participants as compared to German participants during the answering phase (Fig. 14). This difference is most probably of cultural origin. In general, looking at the face of an interaction partner might have two-fold meaning. On the one hand, gazing is an act of perception and attention, as it is needed for experiencing the environment and, particularly, for reading signals from an interaction partner. On the other hand, gaze can be considered a social act in itself, wherein it is used as a signal towards an interaction partner, which has been proposed to be facilitated due to humans' unique (among primates) white sclera. (Argyle & Kendon, 1967; Cook, 1977; Gobel et al., 2015; Kobayashi & Kohshima, 2001). Thus, previous studies have demonstrated that gaze can mean, for example, appealing (Kidwell, 2009), agreeing just as well as disagreeing (Haddington, 2006), and – especially interesting in the context of this study – higher amounts of looking in the eyes while speaking may signal dominance with respect to an interaction partner (Dovidio & Ellyson, 1982, 1985; Hall, Coats, & LeBeau, 2005). Higher level of dominance may be negatively perceived in some cultures, which may be a reasonable explanation for the cultural differences found in this study. Specifically in Japanese culture, it has been shown that displaying dominance is not as positively reinforced as in "Western cultures" (Freeman, Rule, Adams Jr, & Ambady, 2009). Thus, it might be plausible to assume such dominant behavior like extensive looking in the eye or face during speaking may be suppressed in Japanese culture. On the other hand, it is well known that in Japanese culture much of an interaction is based on context and nonverbal cues (Maynard, 1987, 1990; McDaniel, 1993; Morsbach, 1973). Thus, for understanding the meaning of an interaction partner's spoken word, such as during the listening or being questioned phase of our experiment, for the Japanese sample looking at the interaction partner's face may have been essential to read possible signals and thus the context, feelings and meaning of the conversation.

However, one could also interpret the results slightly differently: looking down while answering a question, that is, looking more at the body than the face of the interaction partner, may be an intrinsic, endogenous behavior of thinking, which was naturally displayed for Japanese participants, but was instead overridden in German participants by the social display rule of having to maintain eye contact as a German social norm (McCarthy et al., 2008).

This study's results also demonstrated a general difference in fixation activity occurring for all three interaction types with most fixation activity occurring during the questioning phase, less during the listening phase and least during the answering phase (Fig. 12). This is most probably an effect of the interactive nature of the experiment. Therein, the questioning phase was shortest, as asking a question is usually quick, the listening phase was medium length and the answering phase was longest, as it was based on participants' response and therefore, to keep the natural flow of conversation, not restricted. Hence, while answering (and to a certain extent also while listening) participants had more opportunity to make blinks and saccades, which didn't count as fixations. This, one might argue, may also explain part of the above discussed results. However, for the present study only the relative difference of fixation activity between the three AOIs, Face vs. White Space vs. Body, or the two AOIs, Face vs. Body/White Space, especially in the context of culture, was relevant. Hence the overall difference in fixations for the three interaction types, should not have influenced the interpretation of our main results, as we mostly looked at the gaze patterns within one interaction type. This is further supported by the evident lack of interaction effect between Interaction Type x Country, which shows that any cultural difference is not modulated by the Interaction Type alone, but has to always be considered in light of the relative difference between AOIs (Fig. 14, Fig. 15).

Importantly, participants of this study were embedded in the context of a real social interaction (see also Chapter 1 – Discussion), which in accordance with second-person neuroscience would lead to participants being more readily emotionally engaged. Emotional engagement as well as interaction seem to be the key ingredients for understanding other people's minds and actions. Thus, they are important variables also for investigating social gaze behavior and may be the reason for discrepancies found between studies of real, ecologically valid stimuli and laboratory studies making use of screen-based static stimuli (Cole et al., 2016; Schilbach et al., 2013). In that regard, interacting with a real person has been shown to elicit less gaze towards the interaction partner than would be expected from research conducted with static image stimuli that found a clear gaze preference for fixating eye-based stimuli (Birmingham, Bischof, & Kingstone, 2009; Macdonald & Tatler, 2018). Furthermore, viewing a real environment rather than detachedly observing the same environment on a screen seems to lead to variations in gaze patterns (Foulsham, Walker, & Kingstone, 2011). This is evidence in favour of the importance of using ecologically valid stimuli when investigating social gaze behaviour. As that was the case also in the present study, one may conclude that the cultural variations of gaze behaviour patterns found herein may indeed be typical for behaviour during similar "real" interactions in the respective culture.

2.4.1 Summary and Conclusion

We explored gaze behavior during various types of interaction, such as listening, being questioned and answering in two different cultures, Germany and Japan. The present study demonstrated how, based on the type of interaction, gaze patterns are culture-specific, dependent on the type of interaction. In particular, while answering, German participants tended to look more to the face of their interaction partner as compared to other AOIs, while Japanese participants were equally likely to look to the face or elsewhere. This may be related to different cultural norms that have to be first understood in order to correctly interpret the gaze patterns and any inferences related to intentions or behavior.

Chapter 3 Trustworthiness and Cooperation

3.1 Introduction

Research Question 1:

When judging an interaction partner for trustworthiness and choosing them for cooperation, do people exhibit different gaze patterns predictive of their choices?

Result 1:

No, people seem to show similar gaze patterns, regardless of their choices.

Research Question 2:

Do people from different countries, Germany and Japan, exhibit different gaze patterns predictive of their choices?

Result 2:

Yes, people from Japan seem to look more at the face (and less at the body) of the interaction partner as compared to people from Germany.

Research Question 2_2:

Does that differ from normal listening behaviour (Chapter 2)?

Result 2_2:

Not in Japan, but in Germany participants didn't seem to look much at the body, when they had no judgment goal in mind and when speech was highly informative.

Research Question 3:

Did participants choose a particular interaction partner more than others?

Result 3:

No, all interaction partner seemed to be chosen with similar frequency.

3.1.1 Why Trust in the First Place?

It has been postulated (Lewis & Weigert, 1985) that *Trust* in itself is a prerequisite for society. The reason being that as social entities humans need to know how to act and react in a variety of situations. However, since we can't always perfectly predict our environment, we need to at least reduce the uncertainty of what the immediate future will bring. Consequently, trust is needed, as with trust we can assume that certain possibilities are unlikely to happen. Essentially the uncertainty of the future is reduced enabling us to respond (Lewis & Weigert, 1985). As Simmel argued:

> *" Without the general trust that people have in each other, society itself would disintegrate for very few relationships are based upon what is known with certainty about another person"*
> (Simmel, 1978, pp. 178-179)

In this manner, it has been reasoned that trust is essential for human interaction as in its absence social conflict will emerge (see review on trust Robbins, 2016).

3.1.2 How Does Trust Work? What Is Trust?

Lewis and Weigert (1985) described trust as consisting of 3 components, a cognitive, an emotional and a behavioural component. They explained that for cognitive trust one (hereafter: the truster) collects information about the to be trusted person (hereafter: trustee). However, the information will be inevitably limited, as collected evidence for a prediction will practically never be complete. Thus, on the basis of this limited information a "cognitive leap" towards trust will be made on the assumption that the other party will make a similar "leap".

Nevertheless, besides the cognitive component there is most often also an emotional component involved in trust. This becomes apparent when considering the consequences of a breach of trust. Usually, when this kind of betrayal occurs for a certain matter not only will one not be trusted again for the matter itself, but the whole relationship will be destroyed. Lastly, the behavioural component is the only visible part of trust, the overt action, one displays when trusting, such as for example cooperation. However, Lewis and Weigert (1985) also pointed out that cooperation without trust does exist, so it is occasionally difficult to infer trust based solely from cooperative behaviour.

3.1.3 Trust Depending on Context

Importantly, the extent of cognitive and emotional components seem to vary as depending on the context, for example a meeting of previous strangers versus a longstanding relationship, a different combination of cognitive and emotional trust will develop (Lewis & Weigert, 1985).

Another such context seems to be the power distribution between the interacting parties, as described by Bonoma (1976), 3 different types of trust may emerge based on the "power system". He explained that in so-called unilateral power systems, one party being weaker than the other, trust seems to be based on the credibility of the trustees' personal influence. In a mixed system on the other hand, with a certain equivalence of power, bargaining will occur. For that purpose, trust will emerge not only due to the personal influence credibility of the people involved, but also due to the compliance with the social norms of the bargaining process. Lastly, Bonoma described the bilateral system. Therein, both parties pursue a joint goal. Thus, for trust to develop there exists an underlying assumption that not only the own concerns are considered but also the other party's concerns are cared for (Bonoma, 1976).

In my experiment, presented in this chapter, a meeting of strangers was arranged in a context where a slight unilateral power distribution may be assumed. Specifically, a selection of trustees was presented as being experts and the trusters, in this case the participants, needed to decide on the credibility of each subsequently presented trustee and make a choice in favour of one of them. As the experts were strangers to the participants, a first meeting was arranged, wherein the experts introduced themselves and needed to make a good impression on the participants. Under this condition it is likely to assume that the emotional component of trust will be small. On the other hand, the cognitive component will most likely be similarly small as there may not be a lot of information participants will glean from such a short self-introduction meeting. Thus, the case of *mundane trust* with a small emotional and cognitive component will most likely be established (Lewis & Weigert, 1985).

3.1.4 Nonverbal Behaviour and Culture

Nevertheless, even mundane trust needs a basis for development. In this case it is the information on the trustee one can extract from the first meeting. Interestingly, it has been suggested that not only verbal but especially nonverbal behaviour is essential information for the formation of trust (Boone & Buck, 2003). It has also been shown that verbal cues are less important for perceived trustworthiness of a potential interaction partner than nonverbal cues (Tsankova et al., 2012). However, when taking into account the abundance of nonverbal cues a human displays, it is unclear which nonverbal behaviours seem to be of particular importance in signalling trustworthiness. Especially, when considering that different cultures seem to display different types and different amounts of nonverbal cues.

For example, German culture, which is considered a low-context culture, seems to place more importance on the explicit words spoken during a conversation, whereas Japanese culture is often called a high-context culture, where explicit wording is less important and the implicit conditions of the situation itself are more important (Hall, 1976; Hall & Hall, 2001). And although Japanese culture is considered a high-context culture, it is nevertheless well-known that particularly in a meeting of strangers plentiful gestures are not necessarily part of a Japanese interaction repertoire.

To answer the question which nonverbal behaviour, being facial expressions or bodily gestures, are important during a first meeting within German and within Japanese culture, I conducted the experiment reported in this chapter. Importantly, to solely investigate the influence of nonverbal behaviour on the selection of an interaction partner, the content of speech was disregarded by participants wearing sound-attenuating head gear.

3.1.5 Visual Decision Making and Cognitive Processes Involved

Thus, the choice with whom participants would decide to cooperate further was solely a visual decision based on the nonverbal signals of the serially introduced experts.
Consequently, eye-tracking was used to measure overt visual attention of participants during the self-introduction of their potential future cooperation partner.

Visual overt attention, i.e. eye movements and fixations, is an important indicator for the evidence involved in the decision process of a visual decision task. In principle, to be able to make a decision, participants need to fixate that feature in the visual environment which they base their decision on (Orquin & Loose, 2013).

In recent years the importance of ecologically valid stimuli for investigating gaze patterns has been emphasized (Cole et al., 2016). Furthermore, it has been shown that gaze behaviour may differ when a real person is interacted with rather than when a static image is observed (Birmingham et al., 2009; Macdonald & Tatler, 2018) or even when a life-sized virtual character is used (Skarratt et al., 2010). Similarly, Schilbach et al. (2013) postulate that different neural processes are activated when observing a screen-based stimulus rather than when interacting with a real person, wherein in the latter case emotional engagement is more easily elicited. Accordingly, the present study, like the experiments of Chapter 1 and 2, made use of real humans presenting themselves as stimuli. Thus, a naturalistic scenario involving emotional engagement of participants should have been simulated.

3.1.6 Aim of Study

The aim of this experiment was to investigate which signals of a potential cooperation partners' nonverbal behaviour determined perceived trustworthiness and the decision process in terms of collaboration. Further, I wanted to examine this within German and Japanese culture. Thus, the goal was to test whether there are differences or similarities pertaining to the selection of future interaction partner based solely on nonverbal behaviour.
For that purpose, I designed a real-life interaction and choice scenario with a selection of 4 (in Japan) or 5 (in Germany) future cooperation partners. These cooperation partners were advertised as experts in an economic game that would follow, who needed to be trustworthy so that participants would be able to win the game with their cooperation. Importantly, participants' only criteria, based on which their choice could be made, was a short and sequential self-introduction presentation of each expert. Most importantly, while experts introduced themselves one after the other, participants could only perceive their nonverbal behaviour as verbal content of speech was filtered by the sound-attenuating headphones.

3.2 Materials and Methods

3.2.1 Participants

24 participants took part in the experiment conducted in Japan (Mean age = 19.79; SD = 1.79; 11 male) at the Waseda University, Department of Psychology. 30 participants took part in the experiment conducted in Germany (Mean age = 28.23; SD = 5.89; 17 male) at the Technical University Munich, Institute of Cognitive Systems.

All participants were included in the questionnaire and choice analysis of the experiment. Due to technical problems during data acquisition and calibration procedure data from 17 participants in Japan (Mean age = 19.52; SD = 1.59; 8 male) and 18 participants in Germany (Mean age = 26.94; SD = 3.93; 13 male) were included in the gaze data analysis of the experiment.

All participants gave written informed consent regarding their participation prior to commencing the experiment. They received either monetary compensation or course credit for their participation. Only behavioural (gaze data) and questionnaire data were collected during the experiment. All data were stored and analysed anonymously in accordance to the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standard procedures of the Department of Psychology, Ludwig-Maximilians-Universität München (LMU).

3.2.2 Procedure

The study consisted of two interaction experiments one conducted in Germany and one in Japan.

Before the experiment, participants filled out the Ten Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann Jr, 2003; Muck, Hell, & Gosling, 2007) and autism-spectrum quotient (AQ) - questionnaires (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001).

Participants were instructed in writing that a selection of potential cooperation partners (experts) for a subsequently played economic game would introduce themselves shortly. Further, they were briefed that they would need to choose one of these experts who would teach them the rules of the game and that only a trustworthy source would be able to lead them to win the economic game. Hence, only then would participants be able to win their money for participation in the experiment. In reality, participants received their participation money already after making their choice as there was no subsequent economic game.

Afterwards, sound-attenuating headphones, earplugs and mobile eye tracker were put on. 1-point calibration (or 3-point calibration, if 1-point calibration failed) of the eyetracker in Germany and 9-point calibration of the eye-tracker in Japan commenced and the gaze data recording started. For the purpose of offline calibration of the German eyetracker, participants were instructed to first look at specific points in the environment before the interaction partner would enter the room.

Every experimental session consisted of dyadic interactions between the participant and each of the experts, 5 in Germany and 4 in Japan. During each dyadic interaction the experts introduced themselves for approximately 30 seconds to the participants, who silently observed and nonverbally reacted to them, for example by head nodding or smiling. During each iteration of self-introduction one participant and one experimenter were in the room.

Participants were seated in front of the experts, who were standing while introducing themselves. After each dyadic interaction the current experts left the room after which the next experts entered and positioned himself in front of the seated participant to start the next iteration of self-introduction.

After observing all self-introductions participants had to fill out a choice questionnaire. This questionnaire asked after participants' chosen interaction partner as well as the order in which the other experts would have been chosen and the reasons for each choice.

3.2.3 Stimuli and Apparatus

Participants' gaze data was collected with a mobile eye-tracker. In Japan the Tobii Glasses Eye Tracker (G30-XII Tobii, Tobii Technology Ltd., Japan) with a scene and gaze camera resolution of 30 Hz was used. In Germany the Eye Tracker Glasses (SMI ETG, SensoMotoric Instruments GmbH, Germany) with a scene camera resolution of 30 Hz and a gaze camera resolution of 60 Hz were used.

Viewed "stimuli" consisted of 9 experimenters in total, 5 in Germany (19 - 25 years, 3 female) and 4 in Japan (21 – 27, two male), who presented themselves as experts and introduced themselves for approximately 30 seconds to the participants. All experimenters wore a black turtleneck shirt and blue jeans. Female experimenters wore their hair tied back and a black hairband and all male experts were clean shaven. Participants introduced themselves by following a script: greeting, name, age, study program, hobbies, an argument that they were experts in the economic game and should be chosen, goodbye.

Participants wore earplugs and sound-attenuating headphones (Peltor X5A, 3M), so that they focused their attention on experts' nonverbal behaviour.

3.2.4 Data Analysis

Gaze data was extracted using BeGaze (SensoMotoric Instruments GmbH, Germany) and Tobii Studio (Tobii Technology Ltd., Japan) respectively.

Dwell time was calculated for each participant and stimulus from the sum of gaze durations inside each Area of Interest (AOI). By dividing mean dwell time by trial time the mean dwell time in % was calculated for each participant, stimulus and AOI.

As stimuli were dynamic, Areas of Interest (AOIs) were similarly dynamic, i.e. based on real world criteria instead of fixed measurements. Thus, the AOI Face, was defined according to the face of the experimenter. Similarly, the AOI Gestures was defined as the moving arms and hands of the experimenter. In contrast, the AOI Body was defined as the body of experimenter besides the moving hands and arms. Lastly, the AOI White Space was defined as the whole visual scene short of the experimenter.

The information on preference order assigned to each stimulus by each participant was extracted from the final choice questionnaire.

3.2.4.1 ANOVA - First Choice vs. Last Choice

A mixed-design Analysis of Variance (ANOVA) was conducted for mean dwell time % with within-subjects factors AOI (Face, Body, Gesture, White Space) and Choice (1^{st} Choice, Last Choice) as well as between-subjects factor Culture (Germany, Japan).
Mauchly's test for sphericity was performed and where significant the Greenhouse-Geisser corrected p-values *p[GG]* were used.
95% confidence intervals were calculated based on the correction method from Morey (2008), which takes within-subjects factors into account.

3.2.4.2 ANOVA – Merged Body, Gesture

One further mixed-design ANOVA similar as above was conducted with the only difference being AOIs for Body and Gesture were merged.
Mauchly's test for sphericity was performed and where significant the Greenhouse-Geisser corrected p-values *p[GG]* were used.
95% confidence intervals were calculated based on the correction method from Morey (2008), which takes within-subjects factors into account.

3.2.4.3 ANOVA – Stimulus

Two within-subjects ANOVAS on mean dwell time (%) for the German sample and the Japanese sample were conducted with Stimulus (1-5 for German sample, 1-4 for Japanese sample) and AOI (Face, Body, Gesture, White Space) as within-subjects factors. Two German participants had to be excluded from the analysis as not sufficient eye-tracking data were recorded.
Mauchly's test for sphericity was performed and where significant the Greenhouse-Geisser corrected p-values *p[GG]* were used.

3.2.4.4 ANOVA – Gender

Two within-subjects ANOVAS on mean dwell time (%) for the German sample and the Japanese sample were conducted with participant gender (male, female) as between-subjects factor as well as experimenter gender (male, female) and AOI (Face, Body, Gesture, White Space) as within-subjects factors.

Mauchly's test for sphericity was performed and where significant the Greenhouse-Geisser corrected p-values *p[GG]* were used.

3.2.4.5 Choice -Questionnaire

To test whether participants choose a particular stimulus preferentially a Kruskal-Wallis rank sum test was performed.

3.2.4.6 AQ – Choice

AQ scores were calculated from the AQ test. Where participants couldn't decide between two options in the test, which would correspond to either a point or no point added to the AQ score, half a point was added to the score instead (for two German participants)

Ordinal regression analysis was performed for the German and Japanese sample on Choice as dependent variable with AQ score and Stimulus as independent variable. One participant had to be excluded from the German sample as for two stimuli choice was inconclusively assigned.

3.2.4.7 Personality - Choice

From the Ten Item Personality Inventory (TIPI), scores of categories extroversion, agreeableness, conscientiousness, emotional stability and openness were extracted for every participant.

Ordinal regression analysis was performed for the German and Japanese sample on Choice as dependent variable with TIPI scores for every category and Stimulus as independent variable. One participant had to be excluded from the German sample as for two stimuli choice was inconclusively assigned.

3.2.4.8 Correlations

To perform pearson correlation analyses between participants TIPI and AQ scores and their gaze data, the dwell time-difference scores for Face-Body, Face-Gesture and Face-White Space had to be calculated. P-values of pearson correlations between AQ and difference scores as well as TIPI scores and difference scores were Bonferroni corrected for multiple comparison.

For the purpose of difference score calculation, the average dwell time % per AOI for each participant was calculated. Then the average Body dwell time % was subtracted from the average

Face dwell time % for each participant to generate the Face-Body dwell time-difference score %. Furthermore, the average Gesture dwell time % as well as the average White Space dwell time % were subtracted from the average Face dwell time % for each participant to generate the Face-Gesture dwell time-difference score % and the Face-White Space dwell time-difference score % respectively.

3.3 Results

To gain insight into the gaze patterns during decision making on future interaction partner, gaze dwell time of participants during stimulus presentation, i.e., potential interaction partner's self-introduction, was analysed.

3.3.1 Descriptive Statistics and Outliers

Outliers were calculated based on mean dwell time % for Japanese and German participants and AOIs ± 2SD. One German and one Japanese participant were excluded from further analysis based on outlies (Fig. 16, Fig. 17).

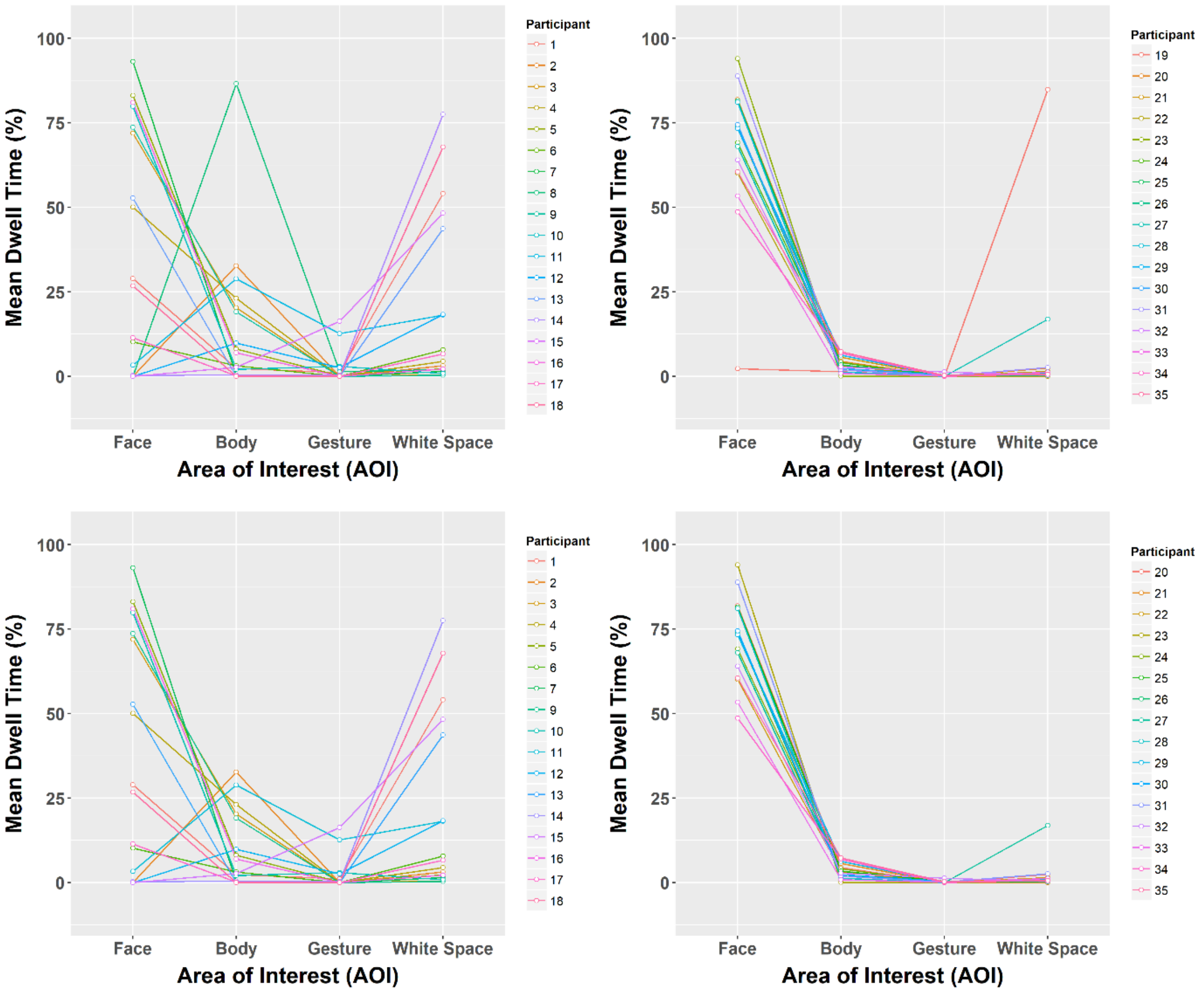

Fig. 16: Mean Dwell Time (%) for all participants and AOIs. Left are German participants (1-18), right are Japanese participants (19-35). Top are with outliers, bottom are without outliers.

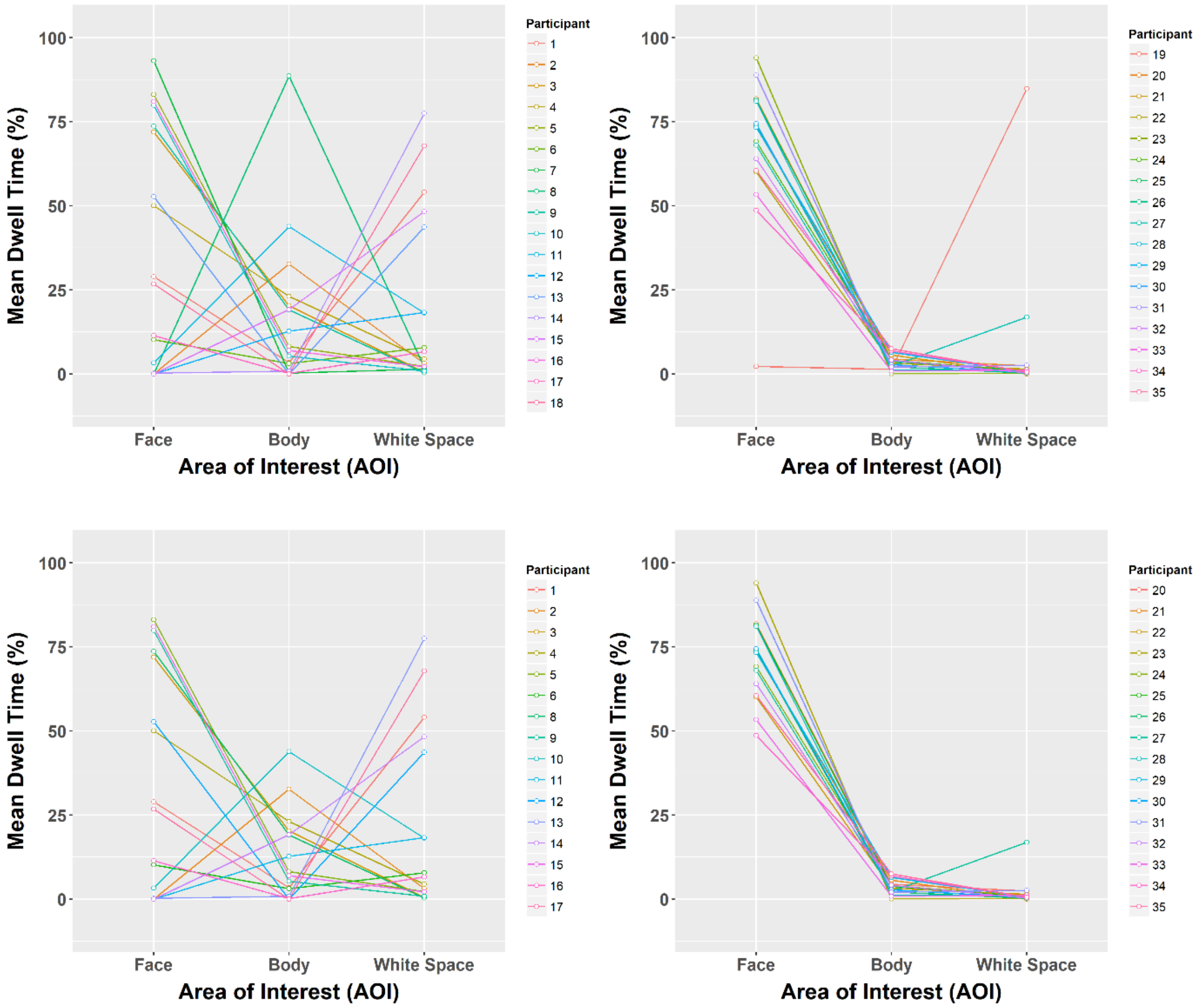

Fig. 17: Mean Dwell Time (%) for all participants and merged AOIs. Left are German participants (1-18), right are Japanese participants (19-35). Top are with outliers, bottom are without outliers.

3.3.2 ANOVA

The mixed-design ANOVA resulted in a significant main effect of Area of Interest as well as an interaction effect between Culture and Area of Interest. No other effect reached the level of significance, $p > .47$.

- Main effect AOI: $F(3, 93) = 58.54$, $\eta^2 = 0.58$, $p[GG] < .001$
- Interaction Effect Culture x AOI: $F(3, 93) = 11.41$, $\eta^2 = 0.21$, $p[GG] < .001$

To investigate the differences between groups in more detail 95% confidence intervals (CI) were compared; where overlapping post-hoc t-tests Bonferroni-corrected for multiple comparisons were performed.

3.3.2.1 Main Effect AOI

As 95% confidence intervals of Body (M = 6.49, 95% CI [2.05, 10.93]) and White Space (M = 11.74, 95% CI [3.02, 20.46]) AOIs were overlapping, a paired t-test was performed and showed equal gaze activity for both AOIs, p = .67 (Fig. 18, left).

However, based on 95% confidence intervals not overlapping and post hoc t-tests, there were differences in activity between all other AOIs, with Face (M = 54.92, 95% CI [43.56, 66.28]) showing highest gaze activity, followed by Body and White Space, who showed equal activity, and lastly the Gesture AOI (M = 1.18, 95% CI [< .001, 3.97]) yielding the least gaze activity (Face>> White Space, Body >Gesture) (Fig. 18, left).
Paired t-tests where confidence intervals overlapped:

- Gesture – Body: t(32) = -3.59, p = .003
- Gesture – White Space: t(32) = -2.97, p = .02

3.3.2.2 Merged Body and Gesture AOI – Main Effect AOI.

Merging Body and Gesture AOI into one AOI Body (see Fig. 18, right) yielded a similar main effect of AOI, F(2, 62) = 45.56, η2= 0.51, p[GG] < .001, with most gaze activity concentrating on the Face (M = 54.92, 95% CI [43.27, 66.57]) as compared to Body (M = 7.77, 95% CI [1.93, 13.61]) and White Space (M = 11.74, 95% CI [2.35, 21.13]) AOI, which showed equal activity according to a post hoc paired t-test, p = .35

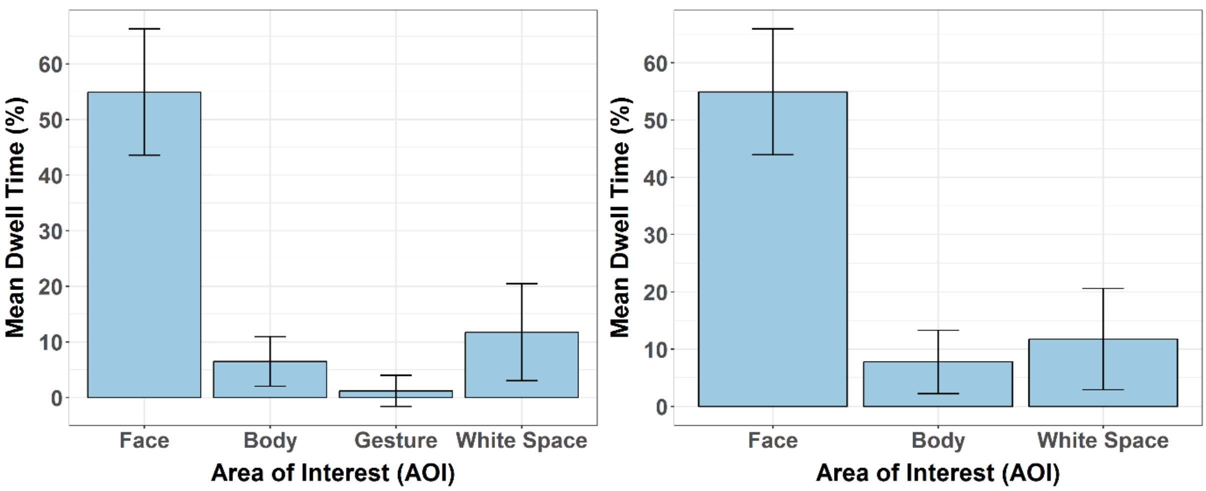

Fig. 18: Mean Dwell Time (%) for AOI (left), as well as merged AOI (right). Error bars: 95% CI

3.3.2.3 Interaction Effect Culture x AOI.

According to 95% confidence intervals not overlapping, German participants showed less gaze activity in the Face AOI than Japanese participants (Germany M = 38.52, 95% CI [20.47, 56.57]; Japan M = 69.80, 95% CI [62.40, 77.20]). By contrast, there seemed to be a trend of German participants showing more gaze activity in the Body AOI than Japanese participants (Germany M = 10.65, 95% CI

[2.20, 19.10]; Japan M = 4.09, 95% CI [0.63, 7.55]), which didn't reach the level of significance according to Bonferroni correction of Welch-Two-Sample-t-test, $p = .35$ with correction, $p = .05$ without correction (Fig. 19, left).

As expected Japanese participants followed the previously established gaze activity of mostly dwelling on the Face AOI, then the Body and White Space (M = 2.19, 95% CI [< .001, 5.21]) which showed equal gaze activity, and lastly fewest gaze activity was shown on the Gesture AOI (M = 0.25, 95% CI [< .001, 2.58]). However, German participants differed from this pattern, as they dwelled equally long on the Face and White Space (M = 22.35, 95% CI [7.18, 37.52]). Nevertheless, also German participants dwelled least on the Gesture AOI (M = 1.24, 95% CI [< .001, 5.72]) and equally long on the White Space and Body.

- Japan: Face>>Body, White Space > Gesture
- Germany: Face ≈ White Space, Body >Gesture

Welch-Two-Sample-t-tests were performed where confidence intervals overlapped:

- Gesture, Germany – Japan: $p > .999$
- Germany, Face – White Space: $p = .92$
- Germany, Body – White Space: $p = .71$
- Germany, Body – Gesture: $t(17.49) = 3.07$, $p = .048$
- Japan, Body – White Space: $p > .999$
- Japan, Body – Gesture: $t(15.26) = 4.05$, $p = .007$

3.3.2.4 Merged Body and Gesture AOI – Interaction Effect Culture x AOI.

According to significant interaction effect of Culture x merged AOI, $F(2, 62) = 119.12$, $\eta^2 = 0.21$, $p[GG] < .001$, and confidence intervals not overlapping German participants seemed to dwell less on the Face (M = 38.52, 95% CI [20.12, 56.92]) and more on the White Space (M = 22.35, 95% CI [6.01, 38.69]) as compared to Japanese participants (Face M = 69.80, 95% CI [62.68, 76.92]; White Space M = 2.19, 95% CI [< .001, 5.94]). Further, German participants tended to also look more at the Body compared to Japanese participants (Germany M = 12.06, 95% CI [1.67, 22.45]; Japan M = 4.33, 95% CI [0.01, 8.65]), according to a Welch-Two-Sample-t-test, $t(18.70) = 2.30$, $p = .03$ without Bonferroni correction (Fig. 19, right).

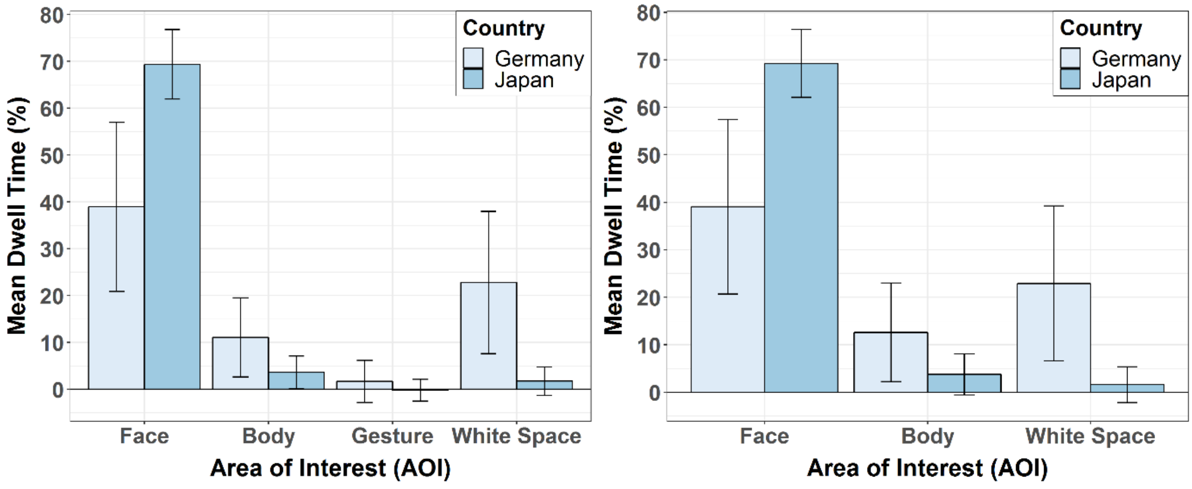

Fig. 19: Mean Dwell Time (%) for AOI and Country (left), as well as merged AOI and Country (right). Error bars: 95% CI

3.3.2.5 Stimulus

To check whether stimulus influenced gaze behaviour, two further within-subjects ANOVAs for the Japanese sample and the German sample separately with AOI and Stimulus as within-subjects factor were conducted.

German sample:

Results showed, similar to the above mixed design ANOVA, only a main effect of AOI (Face, Body, Gesture, White Space), $F(3, 42) = 8.57$, $\eta^2 = 0.31$, $p[GG] < .001$. However, no significant main or interaction effect including Stimulus was found, all $p > .13$.

Japanese sample:

Similarly, only a main effect of AOI, $F(3, 45) = 255.85$, $\eta^2 = 0.90$, $p[GG] < .001$, but no interaction effect of AOI and Stimulus, $p[GG] = .11$, or a main effect of Stimulus, $p = .14$, was found for the Japanese sample.

3.3.2.6 Gender

Neither of the two ANOVAs testing for gender effects found any significant main or interaction effects for experiment or participant gender, all p and $p[GG] > .09$.

3.3.2.7 Choice Influence

As no other main effect or interaction effect reached the level of significance, all $p > .47$, Choice (First Choice vs. Last Choice) also didn't seem to influence gaze behaviour of participants. This was further confirmed with a separate ANOVA which included all choice conditions. (Choice 1, Choice 2, Choice 3/ and Choice 4 (in Germany[1]), Last Choice), all $p > .33$.

[1] As German participants had one stimulus more than Japanese participants, the Choice 3 and Choice 4 allocation of stimuli was treated as equal in the German experiment. To include all Choice options 3 more

3.3.3 Questionnaire Analysis

3.3.3.1 Choice Distribution

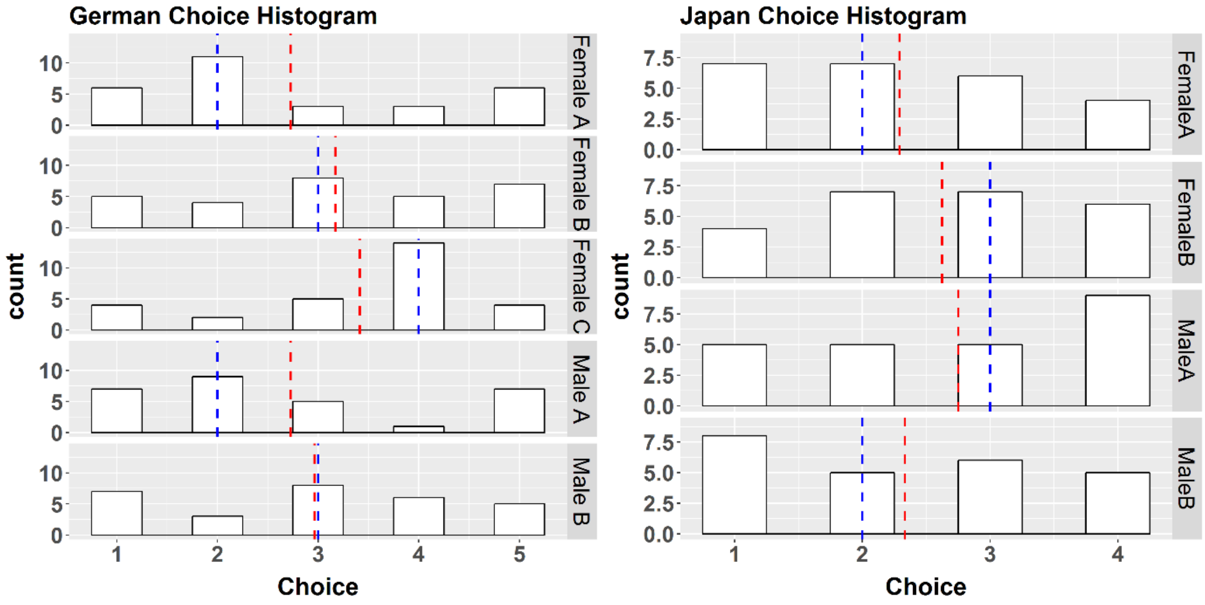

Fig. 20: Choice Histogram for every stimulus (Female A, Female B, (Female C), Male A, Male B) in Germany (left) and in Japan (right). Red dashed line indicates the mean, blue dashed line indicates the median.

The Kruskal-Wallis rank sum test yielded no significant difference between stimuli for neither the German sample, p = .28 nor for the Japanese sample, p = .42.

Thus, overall participants didn't have a clearly defined preferred choice, but all potential interaction partner were equally likely to be chosen by participants (Fig. 20).

---

participants had to be excluded from the German data set, as one participant couldn't decide between assigning 2nd and 3rd choice and two participants didn't have enough eye-tracking data for comparison.

3.3.3.2 AQ Score and Choice

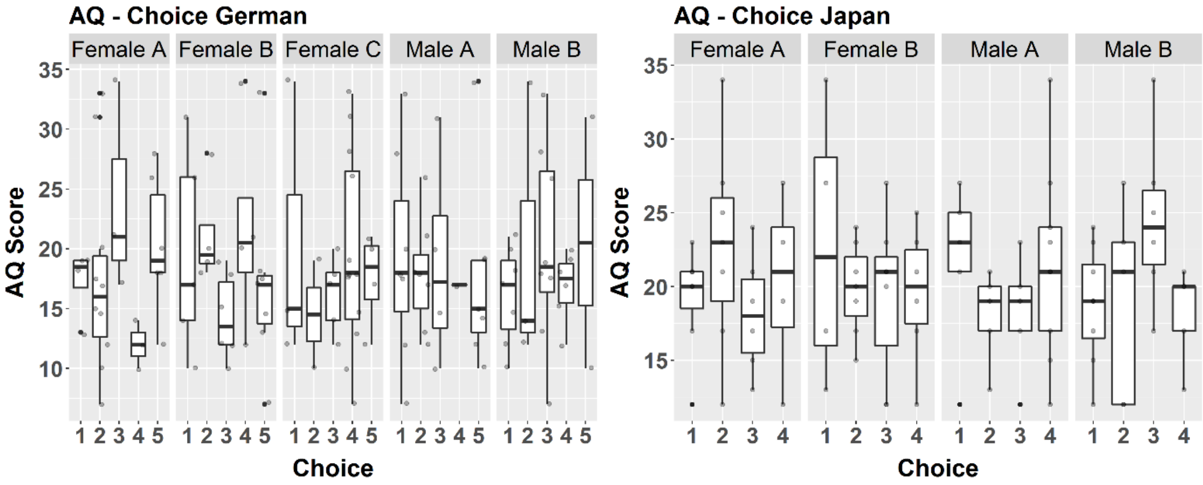

Fig. 21: AQ score distribution of participants for every stimulus and choice in Germany (left) and in Japan (right). Grey and black dots indicate participants (grey inside the whiskers, black outside the whiskers).

To determine whether participants' AQ scores influenced the choice of stimuli an ordinal logistic regression analysis was performed (Fig. 21).

Ordinal logistic regression analysis showed that AQ scores didn't influence choice of stimuli significantly, for neither the German[2] nor the Japanese participants, all $p > .87$.

[2] As German Stimuli violated the proportional odds assumption of ordinal logistic regression analysis, a partial proportional odds ordinal regression analysis was performed for the German sample with Stimulus as nominal factor.

3.3.3.3 Personality and Choice

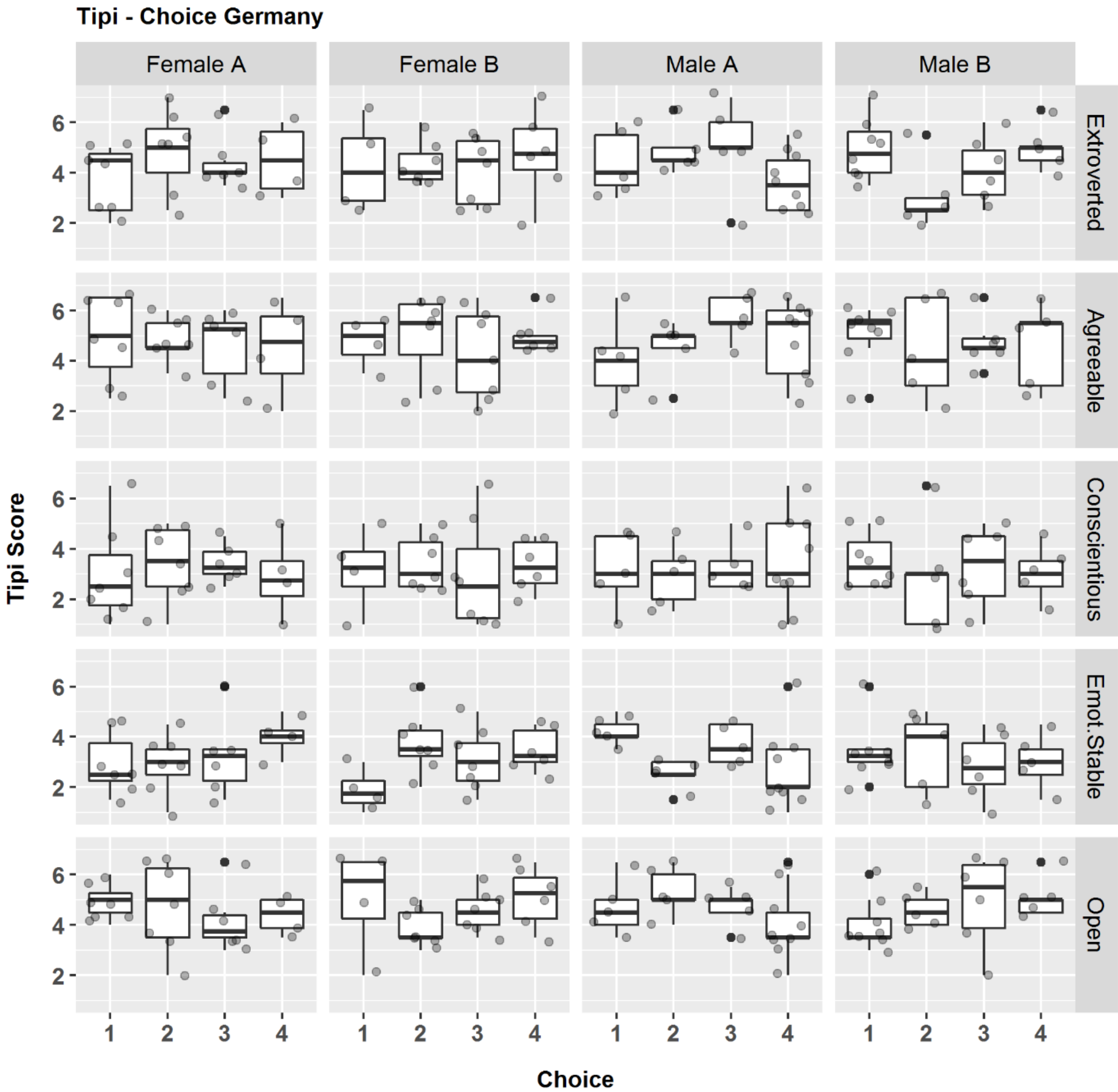

Fig. 22: Ten Item Personality Inventory (TIPI) Score of German participants' extroversion, agreeableness, conscientiousness emotional stability and openness by Stimulus and Choice. Grey dots (inside whiskers) and black dots (outside whiskers) indicate participants.

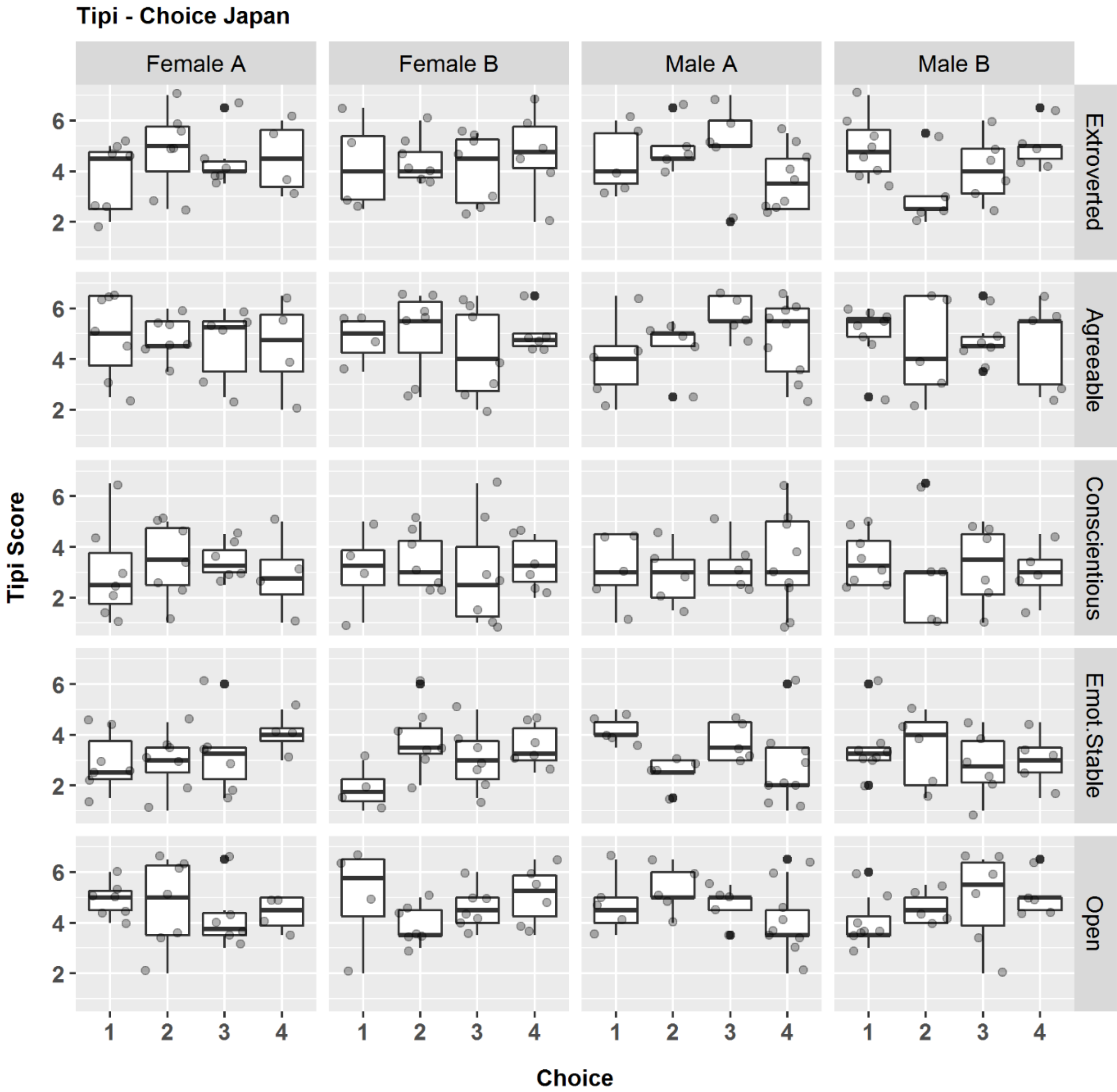

Fig. 23: TIPI Score of Japanese participants' extroversion, agreeableness, conscientiousness, emotional stability and openness by stimulus and choice. Grey dots (inside whiskers) and black dots (outside whiskers) indicate participants.

To determine whether participants' Ten Item Personality Inventory (TIPI) scores influenced their choice of stimuli an ordinal logistic regression analysis was performed. Ordinal logistic regression analysis showed that TIPI scores didn't influence choice of stimuli significantly, for neither the German[3] nor the Japanese participants, all $p > .65$.

[3] As German Stimuli violated the proportional odds assumption of ordinal logistic regression analysis, a partial proportional odds ordinal regression analysis was performed for the German sample with Stimulus as nominal factor.

3.3.3.4 Correlations

No correlation between the dwell time-difference scores (Face-Body, Face-Gestures, Face-White Space, bonferrroni corrected for multiple comparison) and the AQ scores or any of the Ten Item Personality Inventory scores reached the level of significance, $p > .54$.

3.4 Discussion

The aim of this experiment was to investigate how decisions about potential cooperation partner are made. Nonverbal behaviours, that is facial expressions or body movements were examined as a function of cultural background. Specifically, differences between Germany, a low context culture with communicative elements mainly based on the explicit use of language itself, and Japan, a high-context culture with many contextual and nonverbally communicated elements (Hall, 1976), were of interest.

3.4.1 Facial Expression vs. Body Motion

Results showed that in general participants' attention seemed to be attracted by the potential interaction partner, the experts, during their self-introduction. Further, attention remained mainly focused on the face rather than on the body movements and gestures (Fig. 18), which is as anticipated based on results from Chapter 1 and Chapter 2. Moreover, it is an expected result as the face is easily the most salient feature of a human and as participants' goal was to start a relationship in the form of future cooperation based on trust with the introduced experts. Also, although we introduced both female and male cooperation partner, gender didn't seem to influence participants gaze behaviour. Similarly, based on correlation analyses neither personality nor social aptitude seemed to have an effect.

More interesting were the differences between German and Japanese participants. Indeed, German participants tended to look almost twice as much at the body and seemed to look only half as much at the face of the interaction partner, while Japanese participants focused on the face almost exclusively (Fig. 19). Moreover, when considering the whole gaze pattern, a general variability in gaze behaviour across cultures seemed to emerge. On the one hand, Japanese participants seemed to display a gaze pattern typical for listening behaviour as has been shown by results of Chapter 2. Namely, Japanese participants looked mostly at the face, less at the body and only little at the rest of the environment. On the other hand, German participants seemed to deviate from this typical listening gaze pattern. Not only did they look less at the face of the interaction partner, they looked so much less at the face that the amount of gaze towards the face seemed to equal the amount of averted gaze that was displayed. This implies quite a change to both, the results from Chapter 2 and the behaviour of Japanese participants in general. Furthermore, although also German participants seemed to still look less at the body of interaction partners than their face, there appeared to be a trend of increased number of fixations falling on the body, at least compared to Japanese participants. These results may be explained by considering the types of culture that were compared in this study as well as the task participants had to accomplish, namely judgment of trustworthiness.

It is commonly accepted that judgment of trustworthiness is often based on nonverbal communication (Boone & Buck, 2003; Tsankova et al., 2012), especially when considering the notion of deception which is easily accomplished when solely based on content of verbal communication. This may be of particular importance when experts introduce themselves either with similar content of speech or nonverbally and when appearance of introduced experts is also as similar as possible. All three of which had been the case in this study. Experts' appearance was controlled for as much as possible with clothing and hair style and content of speech was not only similar, but also not available to participants, who wore sound attenuating headphones. Thus, the only remaining way to judge trustworthiness was the nonverbal behaviour of the introduced experts.

Moreover, Japan has been conceptualised as a high-context culture, i.e. much information is communicated implicitly or nonverbally based on context. Germany in comparison, may be seen as a low-context culture of much more explicit information transference (Hall, 1976).
However, it also needs to be taken into consideration that Japanese culture is more restrained with respect to body language when communicating with strangers (McDaniel, 1993). Taken together, although Japan may be a high-context culture, i.e., a culture relying on much nonverbal context to communicate information, body motion is much less expressive. Hence, one may conclude that most of the nonverbal communicative context is provided by the face, which may be the reason why participants mostly focused their attention on the face to judge trustworthiness (Fig. 19). In turn, it may be for this reason that Japanese participants showed the same typical "listening gaze pattern" – Face >> Body, White Space – also in a context where they could not listen per se but were only able to rely on presented nonverbal behaviour. In general, it may thus be the case that people from Japan must depend on facial expressions and may not rely on body motion, in both situations when speech is available and when it is not available, as speech alone appears to be mostly non-informative for communication (see Chapter 2) and for social judgment.

German culture, by contrast, commonly uses much more explicit information to communicate (Hall, 1976). Thus, in comparison to Japanese culture people of German culture may rely more on the actual content of speech to communicate rather than the nonverbal behaviour to interpret the speech. However, in the absence of verbal information, as has been the case in this experiment, or when speech becomes otherwise non-informative, nonverbal information may gain more significance and body information in particular may become of more use to gain insight into trustworthiness. Interestingly, in the same manner that speech becomes less informative, German participants seemed to also rely less on the face to gain information, which may have been because the importance of body language tended to be increased.

Thus, the differences in culture-dependant communication style in combination with the specific context of social judgment based on nonverbal behaviour may have been the reason for the cultural differences in gaze patterns between German and Japanese participants.

Importantly, similar to the study of Chapter 2, where gaze behaviour of general communication was examined, also this study presented real people as experts, who acted as stimuli. In this way, it was assured that all recorded gaze behaviour of participants occurred in reaction to being immersed in a real interaction. This may be a crucial distinction, as previous studies have found differences in gaze behaviour of participants who interacted with the real world and participants who observed screen-based laboratory stimuli (Cole et al., 2016; Foulsham, Walker, & Kingstone, 2011; Schilbach et al., 2013). It has been hypothesized that contrary to merely observing laboratory stimuli on a screen an interaction with someone real leads to emotional engagement with that person, which is a pre-requisite for understanding how other people's minds work (Schilbach et al., 2013). Thus, to make inferences about social gaze behaviour it may be of advantage, as was done here, to simulate a natural social environment and let participants interact with real humans.

3.4.2 Choice Behaviour

Interestingly, regarding the choice participants made, no pattern could be established, as not one specific person was chosen more or less often than others, which shows high similarity between experts (Fig. 20). Similar to participants' homogeneous choice preference, participants also didn't seem to differ in their gaze behaviour based on preference of choice stimuli. Thus, one may conclude gaze behaviour during decision making process does not determine choice.

Moreover, neither social aptitude, as based on the AQ questionnaire, nor personality, as based on the Ten Item Personality Inventory, seemed to influence choice of participants in a particular direction (Fig. 21, Fig. 22, Fig. 23). Thus, underlying reasons for a choice and trustworthiness attribution must be either subtler to detect than can be determined by the administered questionnaires or, as this study was based on real interaction, small differences in expert performance might have influenced choice in a way we could not control for.

To test for this, additional experiments containing video stimuli of the experts' introduction instead of real interaction may be considered, even though much communicative content may be lost with loss of embodiment. Although this is a matter of embodiment-control trade-off, it may nevertheless be interesting to consider both sides for comparison. Thus, amongst others, the aim of Chapter 4 has been to use controlled video introduction stimuli as basis for participants' choice behaviour.

3.4.3 Summary and Conclusion

In summary, trustworthiness signals seem to be mainly based on facial expressions, especially for Japanese participants while for German participants' implicit trustworthiness signals seem to be communicated also through body movements.

This is particularly interesting in context of the previous experiment of Chapter 2, which showed that German participants while listening to an interaction partner almost exclusively focused their attention on the face of participants in non-judgment tasks, although it must be taken into consideration that content of speech had been available and of value in that experiment. Nevertheless, one may argue that for German participants first facial expressions and content of speech and next bodily movements are important for communication. Also, when judgments must be based on nonverbal signals alone, either in absence of speech or, more naturally, when content of speech becomes uninformative with respect to choice, bodily movement may be become more prioritised. On the other hand, for Japanese participants independent of judgment or non-judgment tasks and independent of presence or absence of verbal information, mainly facial expressions seem to be the focus of attention, which may be due to the higher communicative content of the face in comparison to the rest of the body.

Chapter 4 Dissecting Nonverbal Behaviour in the Context of Trustworthiness and Cooperation

4.1 Introduction

Research Question 1:

How important is facial expression for judgment of interaction partner? Do participants choose the same person for cooperation when only the face information is available as they choose when viewing the whole body?

Result 1:

Participants choose the same partner only in approximately 30% of the cases, thus body motion seems to provide a lot of crucial information for decision making.

Research Question2:

Does a cultural bias for in-group body movements exist when choosing an interaction partner?

Answer2:

Yes, results showed a bias for cooperating partner with body language from their own culture.

Research Question3:

Do people prefer a particular person or group as their interaction partner and what nonverbal behavior do the preferred interaction partner display?

Answer 3:

Yes, some cooperation partner seemed to be more favored compared to others. In Japan they tended to show high amount of gaze contact and possibly less fidgeting behaviour. In Germany, results are less conclusive, but preferred partner tended to show a mix of moderate amounts of gaze contact and averted gaze, less self-touch and slightly more ideational gestures.

4.1.1 The Face

The face is of particular interest to researchers, who study trustworthiness. The reason becomes clear, when considering the highly developed neural mechanisms underlying face perception (Haxby, Hoffman, & Gobbini, 2000, 2002). Haxby and colleagues (2000, p.223) said:

> *"Face perception, may be the most developed visual perceptual skill in humans."*

Furthermore, the authors suggested that perceiving changes in facial expressions is critical for inferring other people's mental states. In this context they showed evidence for separate perception mechanisms and anatomic structures of dynamic face areas, responsible for facial expression, versus invariant facial areas, responsible for facial identity. Regarding the research community of facial trustworthiness one can perhaps distinguish between two main approaches. The study of physiognomics, which uses static images of faces and in contrast to it the study of moving facial stimuli. I don't differentiate these research streams by facial appearance versus emotional expression, because even in the study of physiognomics one separates the face into static, identity cues and dynamic, emotional cues. Indeed, the latter cues, which hint at emotional expressions even in neutral faces, are the main research focus of trustworthiness judgment in physiognomics.

4.1.1.1 Static Face Stimuli

In that sense, studies on physiognomics showed that judgments on trustworthiness based on static pictures of faces may happen in only 33ms to 100ms (Todorov, 2008; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Todorov et al. (2009) even suggested a saturation point of 167ms after which no changes in judgment will occur anymore. These are judgments solely based on facial geometry of neutral faces. Nevertheless, evidence suggests that similarity of facial geometry to subtle positive or negative emotional expressions is the basis of these trustworthiness judgments. Thus, underlying perception mechanisms of emotional expression seem to be involved in face evaluation. These valence judgments may then be responsible for trustworthiness judgments, because they give information about the harmful versus harmless intent of others. This in turn influences decisions on whether to approach or avoid a person, which is advantageous also from an evolutionary perspective (Oosterhof & Todorov, 2008, 2009; Schaller, 2008). Extensive neuroscientific evidence showed the involvement of the amygdala in emotional expression evaluation (Adolphs, 2010; Sato et al., 2002). Recently, findings on trustworthiness judgments and trait judgments implied to be involved in perceived trustworthiness showed similar results (Adolphs, 2002; Engell, Haxby, & Todorov, 2007; Gordon & Platek, 2009; Todorov, Baron, & Oosterhof, 2008), which corroborates the evidence on the link between emotional expression and trustworthiness judgment. Todorov (2008) even proposes that the high correlation between perceived trustworthiness and other trait judgments, particularly attractiveness, may be a hint that judgment of trustworthiness combines all other trait judgments in itself as an evaluation of valence. Regarding the specific features of facial geometry that participants tend to perceive as more trustworthy, researchers found that high eyebrows, pronounced cheekbones and a shallow nose sellion (Todorov, 2008) as well as narrower facial width (Stirrat & Perrett, 2010) may be involved. However, concerning actual, i.e. correctly judged as measured by a following cooperative task, trustworthiness results remain ambiguous as various studies reported either positive (Stirrat & Perrett, 2010; Tognetti, Berticat, Raymond, & Faurie, 2013), negative or no results (Oosterhof & Todorov, 2008; Todorov, 2008; Tognetti et al., 2013). This shows that perceived trustworthiness and actual trustworthiness at least based solely on facial geometry don't seem to correspond and it has been suggested that any positive results appear to be minimal or may be based on social expectation (Oosterhof & Todorov, 2008; Todorov, 2008).

4.1.1.2 Dynamic Face Stimuli

Another stream of research in the field of face perception is based on moving stimuli of actual emotional expression. In general, emotional expressivity has been suggested to be a signal for actual trustworthiness (Schug, Matsumoto, Horita, Yamagishi, & Bonnet, 2010). From an evolutionary standpoint, this seems reasonable, as emotionally expressive people may have difficulty hiding deceit and thus would be more obliged to cooperate instead (Boone & Buck, 2003; Schug et al., 2010). Interestingly, this expressivity goes in both directions of the spectrum, appropriate positive and negative emotion seem to indicate trustworthiness. Thus, emotional expressions of trustworthiness may be categorized as context dependent, wherein a pro-social emotional expression may be considered trustworthy and a pro-self emotional expression may indicate untrustworthiness (Schug et al., 2010). However, this approach considers actual trustworthiness. What about perceived trustworthiness? Lucas and colleagues (2016) suggested that attribution of trustworthiness is heavily based on facial expressions to the detriment of actual trustworthiness. They found that a more comprehensive attribution including, amongst others, head pose would instead be of advantage. Other studies showed that the subtle differences in facial dynamics of a smile are something humans are able to detect and base their judgment on (Krumhuber et al., 2007). Thus, it seems we are primed to make judgments on facial expressions, not surprising when taking into account the huge perceptual machinery specific to facial identity and facial emotional expression (Haxby et al., 2000, 2002).

4.1.2 The Body

On the other hand, emotional expressivity and valence signals, either positive or negative, may be not only expressed through the face, but also through the rest of the body. Indeed, it has been found that very strong emotions are not signaled through the face, but the body instead as extreme positive and negative facial expressions appear to be indistinguishable from each other (Aviezer et al., 2012). Also, although facial micro expressions are something we usually can't control, in general, bigger facial expressions, which are also easier signals for us to detect, are very well controllable. Deceit is thus also found in smiling persons, similarly as perceived trustworthiness is often judged based on positive emotional valence, of which smiling is a part of. On the other hand, body signals are less well controlled and often considered "leaky" (Argyle, 1988; Ekman & Friesen, 1969a). Along this line, it has been shown that viewing head and body of a person in separate videos influenced trait judgment (Ekman & Friesen, 1969a).

Nevertheless, not much research has been conducted on trust and body perception as compared to the huge corpora of research on trust and face perception. However, if emotional expressivity is a factor when judging trustworthiness, body motion should be considered as a valuable source. After all, it has been shown that humans seem to be very well able to detect emotions from body motions. In fact, there exists a big field of study on action perception and emotion detection. For that purpose, researchers make use of point-light displays to reduce the influence of appearance (Atkinson, Dittrich, Gemmell, & Young, 2004; Chouchourelou, Matsuka, Harber, & Shiffrar, 2006; Dael, Mortillaro, & Scherer, 2012; De Gelder, De Borst, & Watson, 2015). Interestingly, humans seem to not only be able to distinguish emotions, but also make inferences on traits solely based on point-light movements (Montepare & Zebrowitz-McArthur, 1988; Thoresen, Vuong, & Atkinson, 2012).

A similar study showed that evaluation of body motion cues of political speakers superimposed on stick figures seem to be predictive of trait judgments on the original videos of speakers (Koppensteiner, Stephan, & Jäschke, 2015). In a later study Koppensteiner and colleagues showed that for political speakers low gesture activity seemed to increase their trustworthiness (Koppensteiner et al., 2016). On the other hand, it has been demonstrated that the common denominator in perceived trustworthiness - indirectly measured by Big Five psychological trait of conscientiousness - based on gait analysis is the expanded use of personal space. Thus, expanded torso as well as limb movements while walking seemed to be judged as more trustworthy. Interestingly, this trait attribution by participants on the observed walkers didn’t correspond to the subjective trait judgment of the walkers, i.e., when walkers judged themselves (Thoresen et al., 2012). Similarly, a higher rate of gesticulation has been shown to be a positive factor also in perceived persuasiveness (Mehrabian & Williams, 1969), which may be an indirect measure of trustworthiness. Perceived persuasiveness and gestures were also examined in more detail by Maricchiolo and colleagues (2009). They found that ideational gestures, which are related to the semantic content of speech, i.e., deictic, metaphoric, iconic, where overall most positively perceived and that self-touch in particular seemed to reduce judged persuasiveness (Maricchiolo et al., 2009).

4.1.3 Culture

Lastly, I would like to mention culture as an important factor that influences nonverbal behavior in general and trustworthiness signals in particular. For a more comprehensive review see Introduction - Culture and Gaze and Chapter 3 – Introduction – Nonverbal behavior and culture. Thus, in the following study I considered also culture an important factor.

4.1.4 Aim of Study

In this study I investigated social judgment of potential interaction partner for future cooperation based on nonverbal cues in more detail. As opposed to the previous chapter (Chapter 3), which was focused on gaze patterns of participants when confronted with preferred or nonpreferred stimuli, as well as cultural differences in gaze patterns, in this study I investigated solely the choice judgment of various stimuli. Although embodiment of stimuli and corresponding emotional engagement, has been assumed to be an important part of interaction research (Schilbach et. al, 2013), this study aimed at examining details of nonverbal behavior pertaining to perceived trustworthiness and cooperation. Therefore, it was essential that all participants were confronted with the same exact nonverbal behavior. For that purpose, participants in Germany and in Japan took part in three experiments where they viewed video stimuli of people, introduced as experts, showing only their head (Experiment 1), only their body (Experiment 2) and both head and body (Experiment 3). The goal for these three experiments was to answer three research questions. First, I aimed at examining the importance of facial expressions by comparing judgments from Experiment 1 (head stimuli) with those from Experiment 3 (head and body stimuli). I hypothesized that although much research has been focused on face perception, body language plays an equally as high or at least a much higher role than anticipated. Thus, the actual perception of facial expression may not be the most crucial factor for making social decisions on interaction partner.

Secondly, I focused on cultural differences of body language in German and Japanese culture. As it has been shown that humans seem to be biased in favor of faces of their own culture (Ratner, Dotsch, Wigboldus, van Knippenberg, & Amodio, 2014), I wanted to find out whether a similar tendency could be observed for body language. Thus, in Experiment 2 I included body stimuli from the respective other culture, and I compared whether participants in Germany and Japan would preferably choose people from their own culture as interaction partner or whether they would choose body stimuli from the other culture.

Lastly, I investigated whether Japanese and German participants tended to choose one particular person or group of people preferably as interaction partner. I also explored what nonverbal behavior these chosen stimuli presented and whether there existed common behaviors.

Although emotional engagement may be much easier elicited with a real person (see Chapters 1-3), also this study was set-up in a way that aimed at maximizing emotional engagement with the experts presented in the videos as far as possible and without relinquishing the control of showing the same nonverbal behavior across participants. For that purpose, all presented experts faced the camera and thusly, observing participants seemed to be talked at and looked at directly. Furthermore, similar to the study of Chapter 3, a reward was offered for participants making the correct social judgment. In this way, this study aimed at driving participants to try to unearth the experts' underlying motives and possibly through this also induce some emotional engagement and mitigate possible embodiment-control tradeoff effects.

4.2 Materials and Methods

4.2.1 Participants

59 participants took part in the experiment conducted in Germany (Mean age = 23.15; SD = 3.31; 16 male) at the Ludwig-Maximilians-Universität München, Department of Psychology. 60 participants took part in the experiment conducted in Japan (Mean age = 20.58; SD = 3.15; 26 male) at the Waseda University, Department of Psychology.

All participants had normal or corrected-to-normal vision and gave written informed consent regarding their participation prior to commencing the experiment. They received either monetary compensation or course credit for their participation. Only behavioural (gaze data) and questionnaire data were collected during the experiment. All data were stored and analysed anonymously in accordance to the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standard procedures of the Department of Psychology, Ludwig-Maximilians-Universität München (LMU).

4.2.2 Stimuli and Apparatus

The experiment was designed in Experiment Builder (SR Research Ltd., Canada).

Participants were presented with a number of subsequently shown video clips without sound of potential future cooperation partner, introduced as experts for an economic game. To control that participants viewed all video stimuli a static Eyelink 1000 eye tracker (SR Research Ltd., Canada) was used and participants' head movements were restricted via chin-head rest. Each video stimulus consisted of one person, the expert, introducing himself for approximately 30 seconds against a green background. All experts were wearing a black turtleneck shirt and blue jeans. Female experts wore their hair tied back and a black hairband. Male experts all had a clean shaved face.

In Experiment 1 only the head of experts was displayed in the videos; 11 stimuli in total consisting of German experts in Germany (5 male, including 4 experts (2 male) from Chapter 3) as well as Japanese experts in Japan (6 male, including 4 experts (2 male) from Chapter 3) were used.

In Experiment 2 only the body from neck to middle of upper thigh were displayed in the videos; 21 experts were used in total, 11 of which were from participants' own culture (same experts as in Experiment 1) and 10 (5 male) of which were of the respective other culture (Japan or German).

In Experiment 3 both the head and the body were displayed in the videos; 11 stimuli in total consisting of German experts in Germany (5 male, same as in Experiment 1) as well as Japanese experts in Japan (6 male, same as in Experiment 1) were used.

4.2.3 Procedure

The study consisted of 6 experiments, Experiment 1 - 3 were conducted in Germany and in Japan respectively. All participants of the German and Japanese sample respectively took part in all three experiments, which were conducted in series.

At the start of every experiment participants were instructed in writing that they would view a selection of video stimuli from which they needed to choose a cooperation partner for an economic game, the person they would feel was most trustworthy. Afterwards calibration commenced during which participants sequentially fixated 9 positions on the screen. Then, the experiment began. Every trial started with the identity number or letter of the following video, shown for 3.9s. Then a black fixation cross appeared on a grey background for 1s, to make participants aware of the start of the video. Subsequently, a blank grey background flashed for 500ms, after which the corresponding video played. Participants were instructed to view the video in full length. Afterwards they had 5s to make a note on a prepared note paper about the viewed video stimulus to help in their judgment later on. During the note taking, the word memo was presented on the screen on a grey background. Participants were instructed to only move their eyes to make the note and the note paper was positioned accordingly. Then the next trial started.

At the end of ever experiment participants filled out a questionnaire to name whom they would choose as interaction partner. They also indicated places 2-5 as well as the last place. German participants also indicated in a separate question, whom they felt was most trustworthy, as a pilot experiment showed that German participants oftentimes didn’t choose the most trustworthy person as cooperation partner.

4.2.4 Data Analysis

Choice information was extracted from the questionnaires. Nonverbal behavior of stimuli was extracted from the video. Before analysis of the experiments every stimulus (Experiment 1, 2, 3) was rated by me on gaze movements. For that purpose, videos were manually rated on number of straight or averted gaze (every blink counted as additional gaze), number of smiles and number of eyebrow movements. Furthermore, for body stimuli number of gestures were analyzed, particularly number of ideational, conversational and ideational+conversational gestures (= Gestures) as well as number of self-touches were counted. Lastly number of body movements such as up and down or left and right movements were counted and summarized as Kinetics. To check whether all participants looked at the screen while video stimuli were running dwell time in % at the area of displayed video stimulus was extracted from EyeLink DataViewer (SR Research Ltd., Canada).

4.2.4.1 Question 1: Comparison Experiment 1 Head and Experiment 3 Head+Body

One of the possible outcomes (potential outcome 1) of research question 1 was that most participants choose the same person in Experiment 1, when only the head of experts was visible, and Experiment 3, when head and body were visible. Ideally, in this case (potential outcome 1) one would be able to conclude that the face region may be most important for social judgment of trustworthiness. However, one would still not be able to exclude that chosen experts showed also highly preferred body language in addition to highly preferred facial expressions. To come closer to the ideal of separating out experts with highly preferred body language and facial expression, I conducted the following analysis: First, the proportion of participants choosing the same person while viewing the Head (Experiment 1) and the Body (Experiment 2) was calculated for German, and Japanese participants and a two-sample test for the equality of proportions was calculated. These participants were removed from further analysis, so that a comparison between participants from Experiment 1 Head with participants from Experiment 3 Head+Body could be made without the confounding element of participants that choose the same Head (Experiment 1) and Body (Experiment 2). Thus, if most participants still choose the same person in Experiments 1 and 3 it may be more likely that experts were chosen based only on the face information. On the other hand, if results showed that only a small number of participants chose the same expert (potential outcome 2) in Experiments 1 and 3, the previous exclusion would have become redundant. Thus, also an additional analysis with the whole set of participants included was calculated for comparison.

Next, the proportion of participants deciding on the same person as their first choice for Experiment 1, i.e., while viewing only the head, and Experiment 3, i.e. while viewing the head and the body, was calculated. Separate calculations for Japanese, German and both German and Japanese participants were conducted.

4.2.4.2 Question 2: Experiment 2 Body, In-Group vs Out-Group Bias

Participants could decide on first, second, third and fourth choice (Choice 1 - 4) out of 21 viewed body stimuli. For every participant the proportion of stimuli chosen from own culture (German, Japanese) was calculated by dividing the number of own culture stimuli for any of the Choices 1 - 4 by the maximum number of choices, four[4]. Subsequently, two One-sample-t-tests (1-sided) were performed over the average proportion for participants in Germany and in Japan and a Welch-Two-Sample-t-test for comparison of Japanese and German data was calculated.

[4] German participants could also indicate whom they trusted most. In cases where this choice didn't correspond with the Choices 1- 4 (N=4), one additional Choice 5 was included. In this case proportions were based on 5 instead of 4 maximum number of choices.

4.2.4.3 Question 3: Preferred Choice and Specific Nonverbal Signals

For every participant stimuli were scored based on participants' choice and the number of stimuli involved. For experiment 1 and 3, the chosen stimuli received the score 1 - 4 respectively[5], the stimulus chosen as last choice received the score 11 and the stimuli that weren't chosen received a score of 8, the average (rounded up) of places 5 - 10. Similarly, for experiment 2, the chosen stimuli received the score 1 - 4 respectively[5], the stimulus chosen as last choice received the score 21 and the stimuli that weren't chosen received a score of 13, the average of places 5 -21.

Afterwards a Kruskal-Wallis rank sum test was performed on choice scores. Subsequently, stimuli were separated into preferred (high choice) and not preferred (low choice) group. Afterwards, nonverbal behaviour of preferred stimuli was explored based on previously extracted video information.

[5] German participants could also indicate whom they trusted most. In cases where this choice didn't correspond with the Choices 1- 4, one additional Choice 5, i.e. score 5, was included.

4.3 Results

4.3.1 Participant Removal

German participants looked at the video stimuli on average 99.90% (SD = 0.33%) of the time in the first experiment, 99.67% (SD = 0.86%) of the time in the second experiment and 99.78% (SD = 0.99%) of the time in the third experiment at the video stimuli. Thus, all participants remained included in the study.

Japanese participants looked on average 97.86% (SD = 5.56%) of the time in the first experiment, 96.89% (SD = 4.79%) of the time in the second experiment and 98.35% (SD = 4.33%) of the time in the third experiment at the video stimuli.

One participant that looked only 60.10 % of the time in the first experiment, one participant that looked only 72.67% of the time in the second experiment and one participant that looked only 71.03% of the time in the third experiment were excluded from the analysis of the respective experiment. Further, one participant was excluded from the third experiment as the first two trials didn't run.

4.3.2 Question 1: Comparison Experiment 1 - Head and Experiment 3 - Head+Body

I was interested in behavior of participants, who choose different experts when viewing only the head and only the body. In Germany 15.25% (N=9) and in Japan 27.27% (N=15) of participants choose the same person while viewing the Head (Experiment 1) and the Body (Experiment2) and were removed from analysis. There remained 50 German and 40 Japanese participants.

Next, these participants' choice on experts while viewing only their head and only their body was analyzed:

The proportion of participants deciding on the same person as their first choice while viewing only the head (Experiment 1) and while viewing the head + body (Experiment 3), amounted to:

- 0.32 for German participants (N=50), 95% CI [0.21, 1.0]
- 0.33 for Japanese participants (N=40), 95% CI [0.20, 1.0]
- 0.32 for combined German and Japanese participants (N=90), 95% CI [0.24, 1.0]

Thus, most participants seemed to choose different experts when seeing both head+body as compared to seeing only the head (Fig. 24).

Interestingly, the same analysis with participants[6] who choose the same person in Experiment 1 (Head) and Experiment 2 (Body) included yielded similar results:

- 0.32 for German participants (N=59), 95% CI [0.22, 1.0]
- 0.35 for Japanese participants (N=55), 95% CI [0.24, 1.0]
- 0.33 for combined German and Japanese participants (N=114), 95% CI [0.26, 1.0]

Again, participants seemed to choose differently when viewing the whole body as compared to viewing only the head.

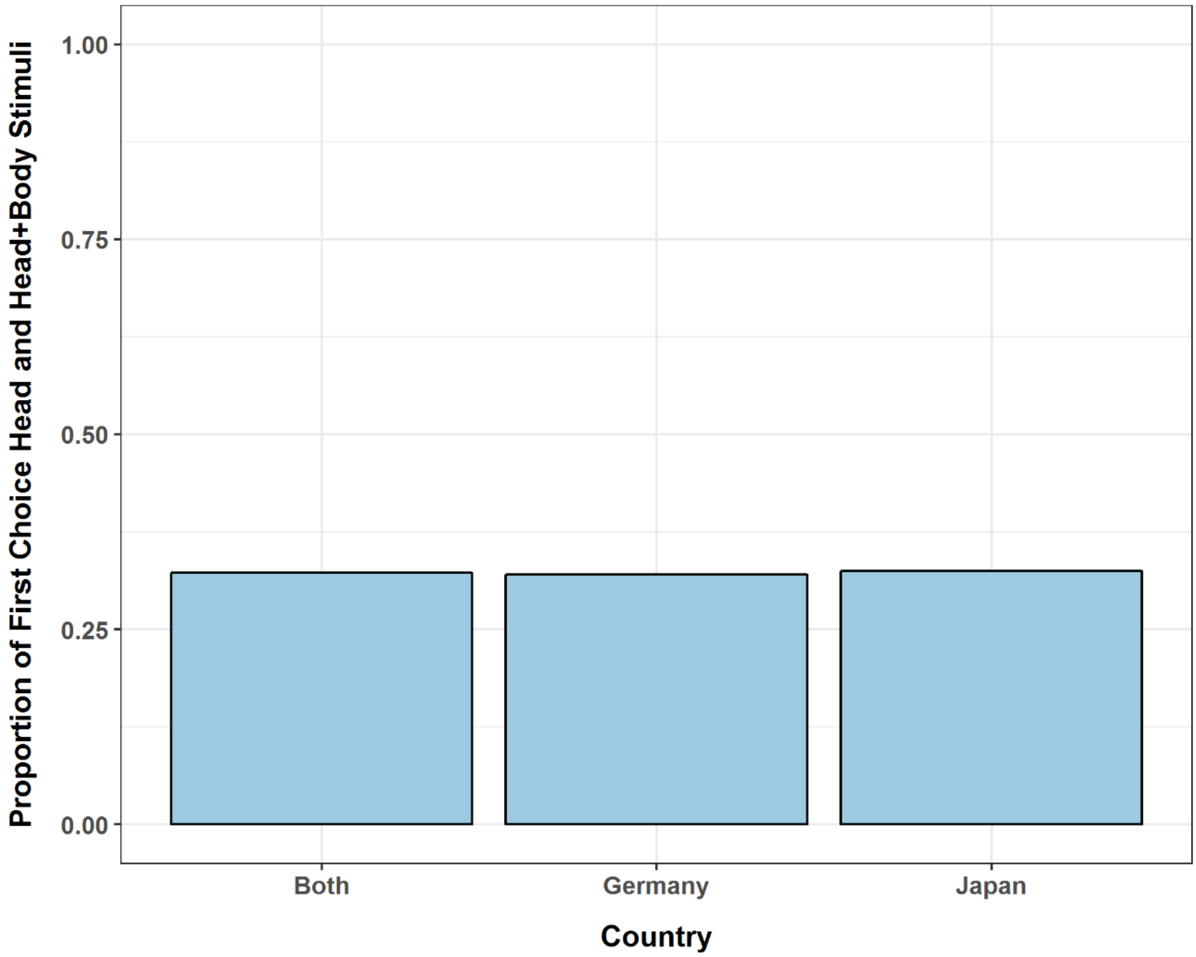

Fig. 24: Proportion of participants choosing the same stimulus as First Choice after viewing the head and after viewing the head and the body.

4.3.3 Question 2: Experiment 2 Body, In-Group vs Out-Group Bias

One additional Japanese participant was removed from analysis of Experiment 2, because Choice 3 wasn't indicated in the questionnaire.

[6] One German participant choose the same person for choice 1 and 3. I included this person in the analysis.

The mean proportion for choosing a body stimulus from their own culture amounted to:

- 0.73 for German participants
- 0.60 for Japanese participants

A One-sample-t-test (1-sided) showed that the average proportion of German participants choosing a body from their own culture, 0.73 (95% CI [0.68, Inf]), was significantly higher than chance (0.5), $t(58) = 22.60$, $p < .001$.

Similarly, a One-sample-t-test (1-sided) yielded a significantly higher than chance average proportion of Japanese participants choosing a body stimulus from their own culture, 0.60 (95% CI [0.54, Inf]), $t(56) = 17.01$, $p < .001$.

A Welch-Two-Sample-t-test showed that a significantly higher proportion of German participants choose a body stimulus from their own culture as compared to Japanese participants, $t(112.86) = 2.79$, $p = .006$.

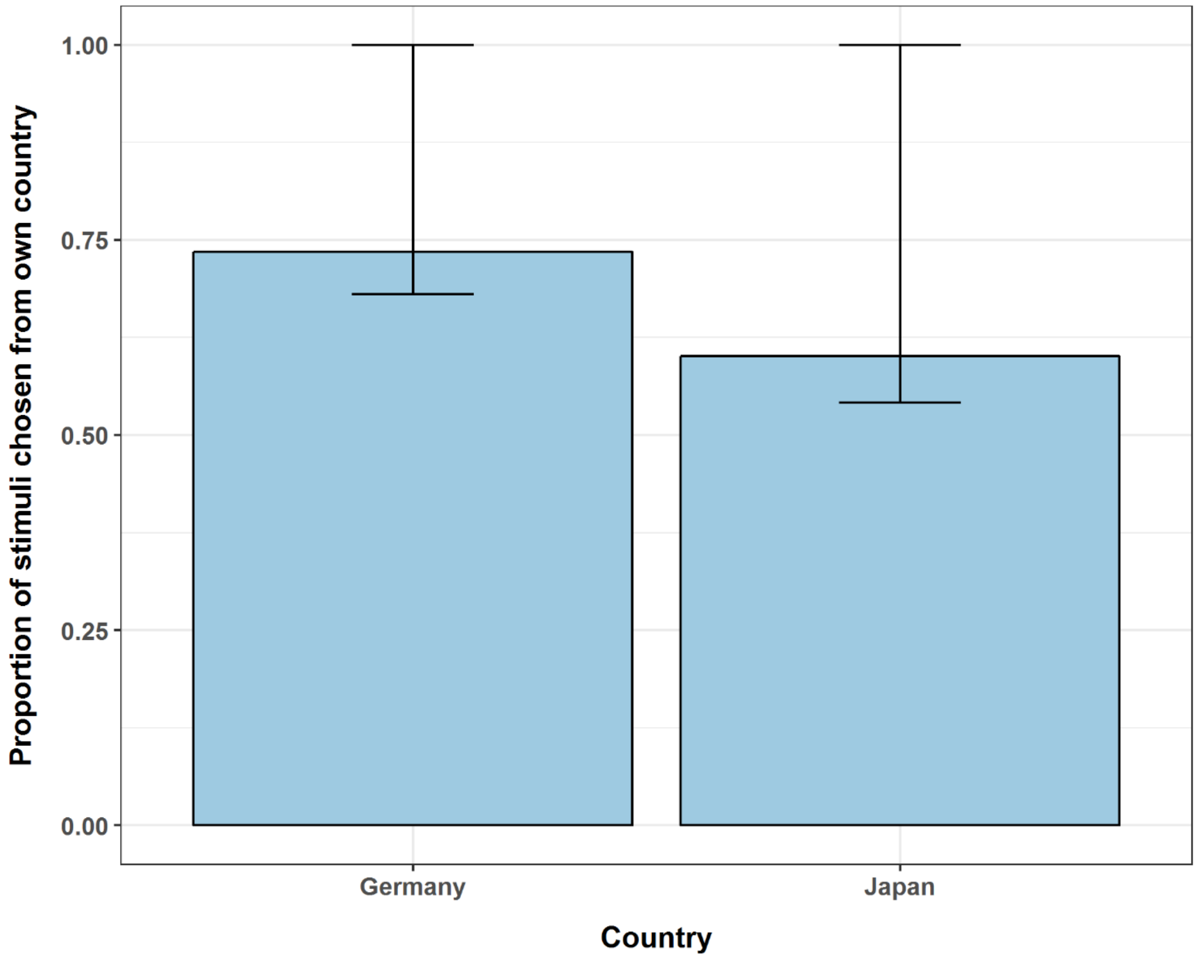

Fig. 25: Average proportion of chosen stimuli from own country. Error bars are 95% CI.

4.3.4 Question 3: Preferred Choice and Specific Nonverbal Signals of Choice

One additional Japanese participant was removed from Experiment 1 analysis as choice 3 and last choice were the same. In experiment 1, one German participant had the same choice 1 and choice 3, in this case I only removed choice 3 from analysis.

The Kruskal-Wallis rank sum test was performed to analyse whether stimuli were rated differently on their choice score. It yielded a significant difference for stimuli of all experiments for German and Japanese participants:

- Experiment 1 Head German: X^2 (10, N=59) = 37.01, $p < .001$
- Experiment 1 Head Japan: X^2 (10, N=57) = 87.24, $p < .001$
- Experiment 2 Body German: X^2 (20, N=59) = 124.64, $p < .001$
- Experiment 2 Body Japan: X^2 (20, N=58) = 128.39, $p < .001$
- Experiment 3 Head+Body German: X^2 (10, N=59) = 29.34, $p = .001$
- Experiment 3 Head+Body Japan: X^2 (10, N=57) = 82.14, $p < .001$

To check which specific stimuli were rated higher on the choice score compared to others a post-hoc Dunn's multiple comparison test with p-values adjusted by the Benjamin-Hochberg method was performed. The results are displayed in Fig. 26 - Fig. 29, same letters indicate no significant difference between stimuli.

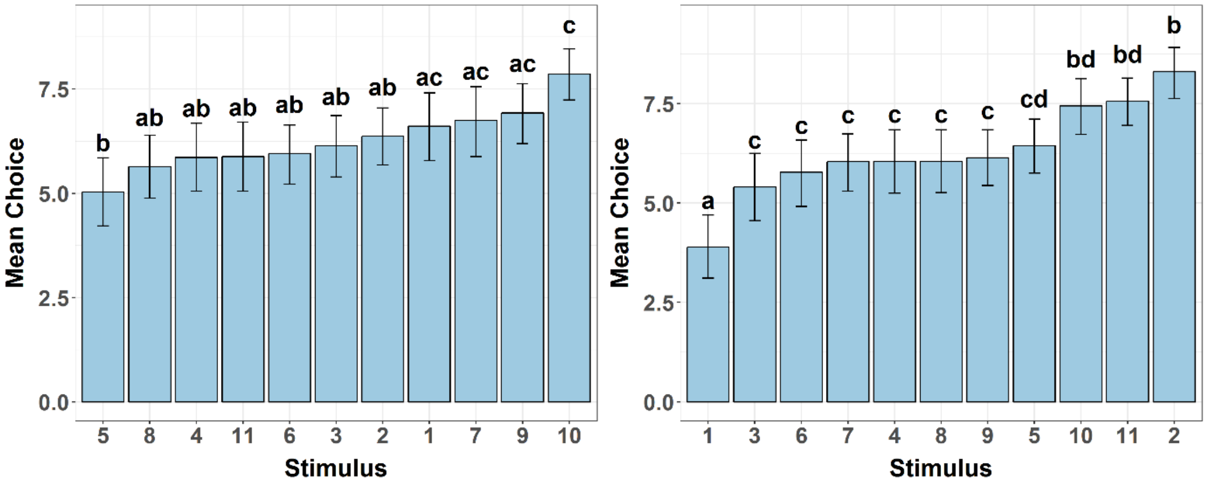

Fig. 26: Experiment 1 - Head. German participants' (left) and Japanese participants' (right) mean choice score with Dunn's multiple comparison test results displayed as letters above error bars, same letter indicate non-significant difference. Low score indicates better choice. Error bars are 95% confidence intervals (Nonparametric Bootstrap Confidence Intervals).

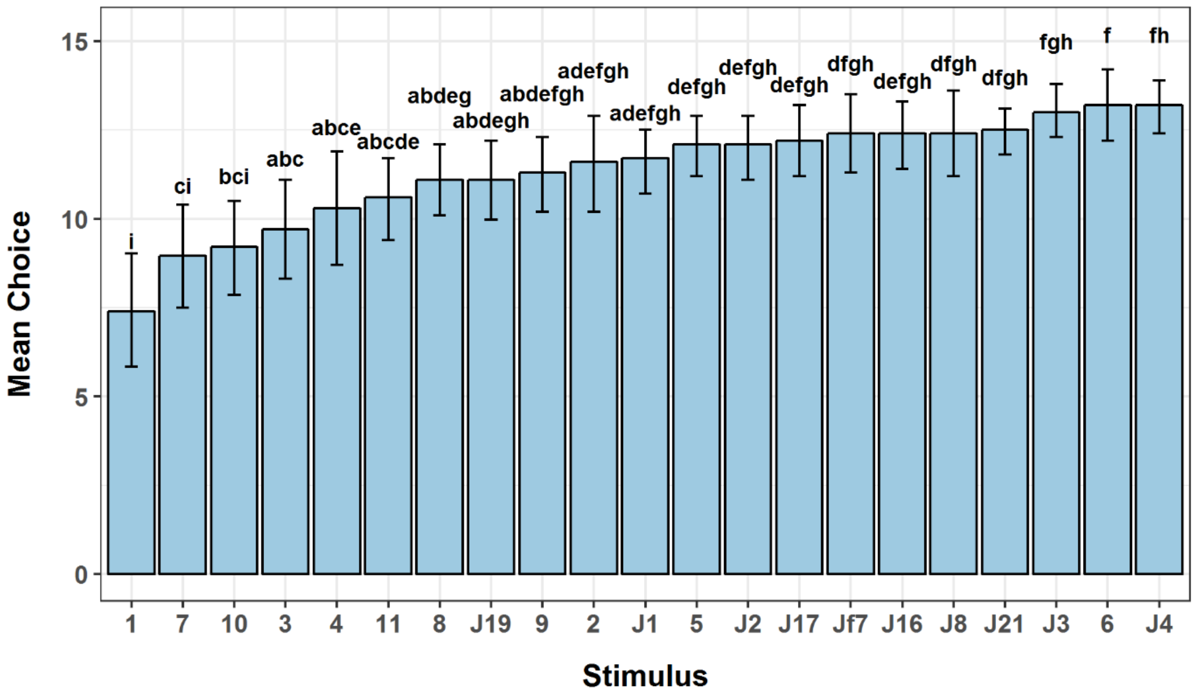

Fig. 27: Experiment 2 - Body. German participants' mean choice score with Dunn's multiple comparison test results displayed as letters above error bars. Same letter indicate non-significant difference. Error bars are 95% confidence intervals (Nonparametric Bootstrap Confidence Intervals).

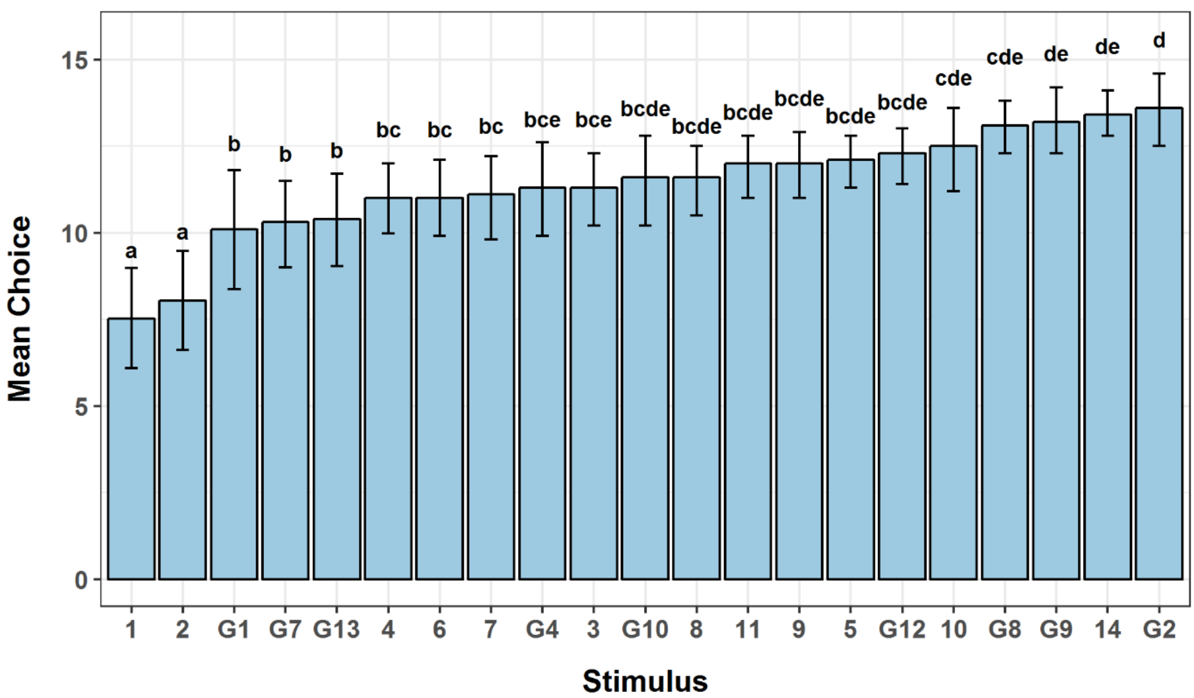

Fig. 28: Experiment 2 - Body. Japanese participants' mean choice score with Dunn's multiple comparison test results displayed as letters above error bars. Same letter indicate non-significant difference. Error bars are 95% confidence intervals (Nonparametric Bootstrap Confidence Intervals).

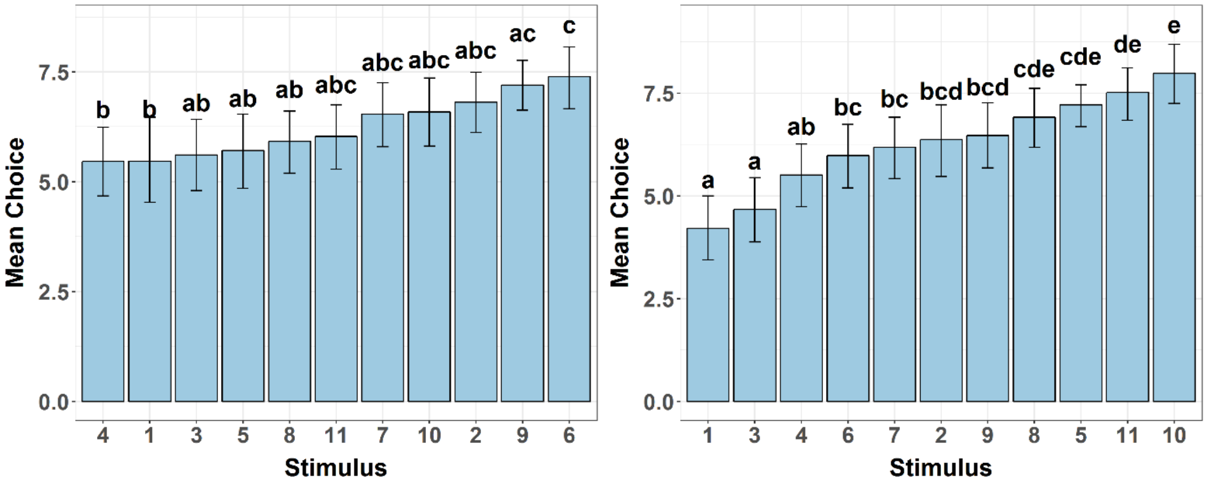

Fig. 29: Experiment 3 - Head+Body. German participants' (left) and Japanese participants (right) mean choice score with Dunn's multiple comparison test results displayed as letters above error bars. Same letter indicates non-significant difference. Error bars are 95% confidence intervals (Nonparametric Bootstrap Confidence Intervals).

Although, only few stimuli were chosen significantly more than others, based on results of Dunn's test I separated stimuli further into high or low, or in rare cases also middle, choice category (see Fig. 26 - Fig. 29). High category stimuli for the different experiments included:

- Experiment 1 - Head German: High category = Stimuli 5, 8, 4, 11, 6, 3, 2
- Experiment 1 - Head Japan: High category = Stimuli 1, 3, 6, 7, 4, 8, 9
- Experiment 2 - Body German: High category = Stimuli 1, 7, 10, 3, 4, 11
- Experiment 2 - Body Japan: High category = Stimuli 1, 2, G1, G7, G13
- Experiment 1 - Head+Body German: High category = Stimuli 4, 1, 3, 5, 8
- Experiment 1 - Head+Body Japan: High category = Stimuli 1, 3, 4

The following figure (Fig. 30) presents the nonverbal behavior displayed by above categorized High and Low Choice head stimuli of Experiment 1.

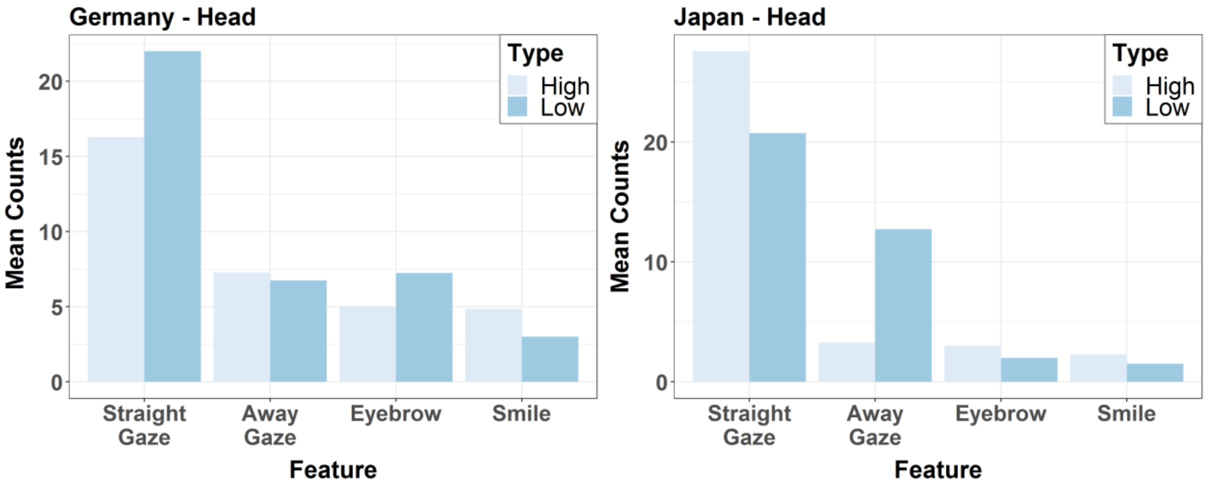

Fig. 30: Head. Mean counts of gazing straight, gazing away, moving eyebrows and smiling for German (left) and Japanese (right) as well as high and low categorized stimuli.

Exploration of nonverbal behaviour of head stimuli demonstrated that preferred German interaction partner tended to show less straight gaze (16.29) and approximately equal amounts of averted gaze (high = 7.29, low = 6.75) compared to less preferred (22.0) interaction partner. Japanese data showed the opposite trend with preferred interaction partner displaying more straight gaze (high = 27.57, low = 20.75) and less away gaze (high = 3.28, low = 12.75). Overall stimuli didn't tend to show big variations in eyebrow movements (Germany high = 5.0 , low = 7.25; Japan high = 3.0, low = 2.0) and amount of smiling (Germany high = 4.86, low = 3.0; Japan high = 2.29, low = 1.5) (Fig. 30).

For the sake of completeness also nonverbal behavior displayed by the separate head stimuli of Experiment 1 ordered by choice according to Dunn's test is presented (see Fig. 31).

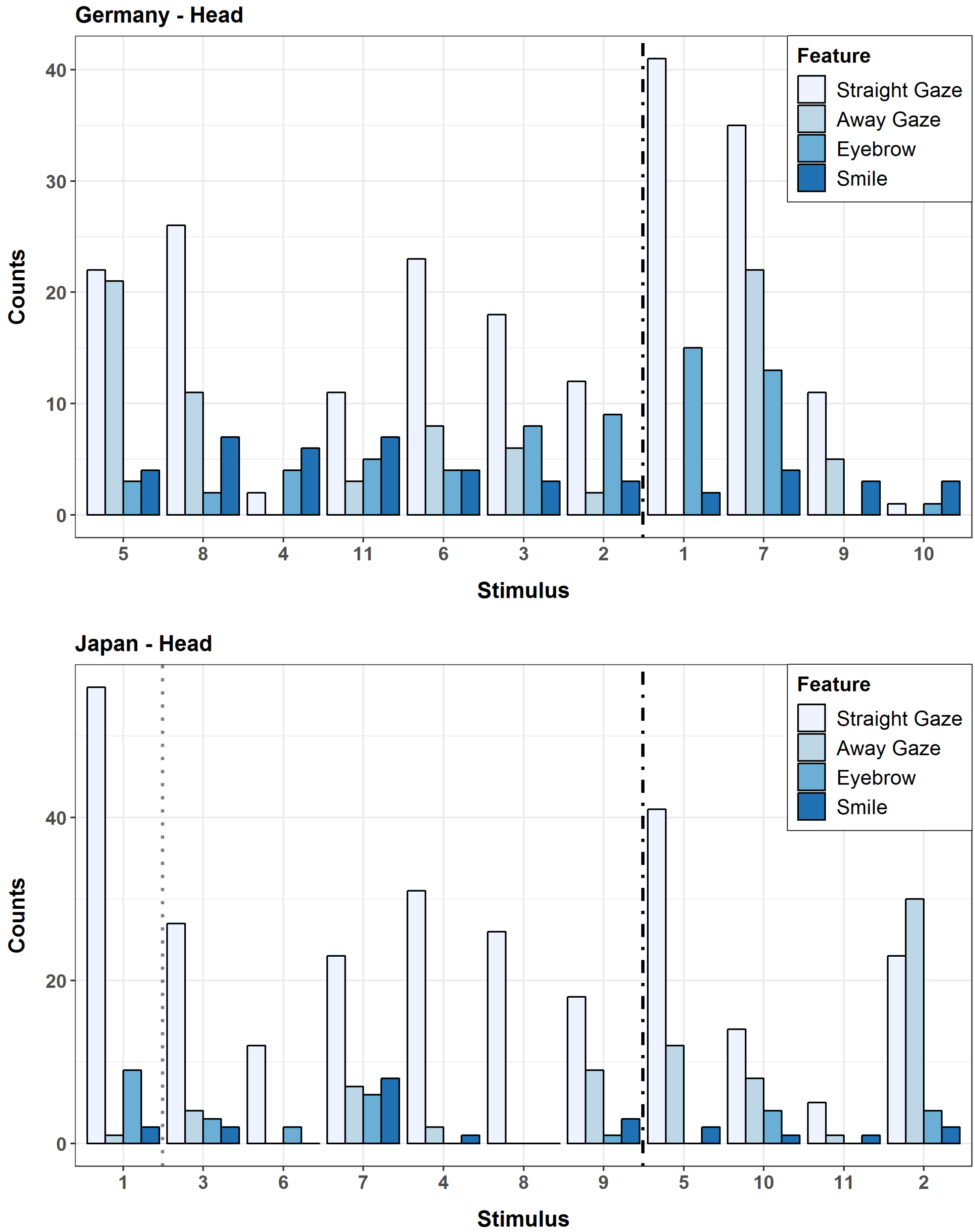

Fig. 31: Head. Counts of gazing straight, gazing away, eyebrow movements and smiling for every stimulus for German (upper) and Japanese (lower) experiment. Dashed line indicates the boundary between high (left) and low (right) categorized stimuli. Dotted line indicates the first stimulus, which had been significantly higher than all other stimuli.

The following figure (Fig. 32) presents the nonverbal behavior displayed by above categorized High, Low and Middle Choice body stimuli of Experiment 2.

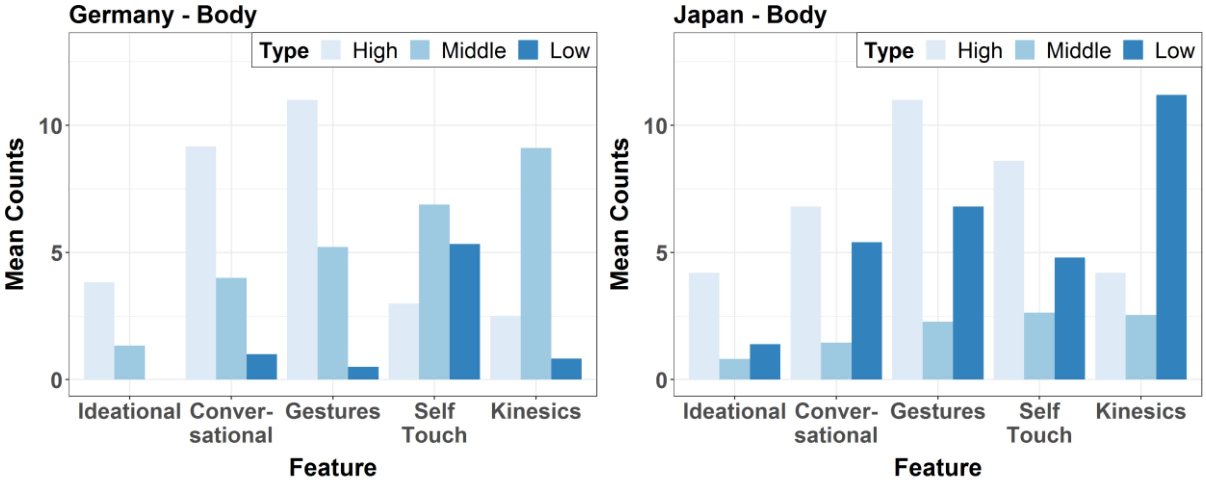

Fig. 32: Body. Mean counts of ideational gestures, conversational gestures, ideational+conversational (Gestures), self-touch and kinesics for German (left) and Japanese (right) experiments and high, middle and low categorized stimuli.

Exploration of nonverbal behaviour of body stimuli revealed that preferred German interaction partner tended to display more gestures (11.0) and slightly less self-touch (3.0) compared to their less preferred counterparts (gestures 0.50, self-touch = 5.33). Results for ideational (high = 3.9, middle = 1.33, low = 0.00) and conversational gestures (high = 9.17, middle = 4.0, low = 1.0) reflected combined results for gestures. Japanese data showed a similar trend with preferred interaction partner displaying more gestures (high = 11, low = 6.8). However, interaction partner from the low group tended to show also more gestures than the middle group (2.27). A similar trend emerged also for self-touch, high (8.60) > low (4.80) > middle (4.63). Ideational (high = 4.20, middle = 0.82, low = 1.40) and conversational gesture counts (high = 6.80, middle = 1.45, low = 5.40) reflected results from combined gestures.

Regarding kinesics, German interaction partner from the middle group tended to display highest amount of body movements (9.11) as compared to the high (2.50) and low group (0.83). Japanese preferred interaction partner on the other hand, tended to show much fewer body movements (4.2) than the less preferred interaction partner, who tended to move much more (11.20) (Fig. 32).

For the sake of completeness also nonverbal behavior displayed by the separate body stimuli of Experiment 2 ordered by choice according to Dunn's test is presented (see Fig. 33).

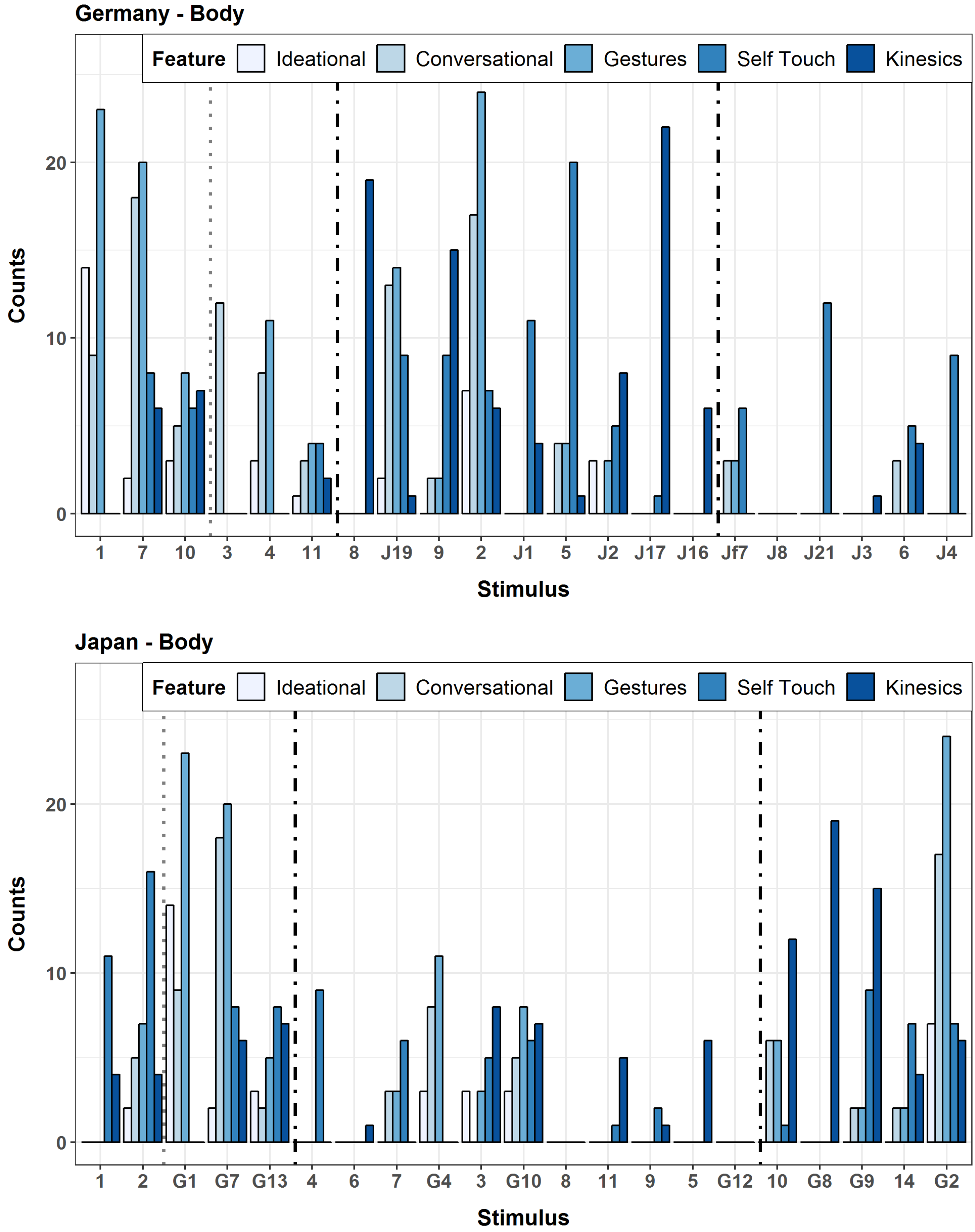

Fig. 33: Body. Counts of ideational gestures, conversational gestures, ideational+conversational (Gestures), self-touch and kinesics for every stimulus for German (upper) and Japanese (lower) experiment. Dashed lines indicate the boundary between high (left), middle (middle) and low (right) categorized stimuli. Dotted line indicates the first stimuli, which had a significantly higher choice score than all other stimuli.

The following figures (Fig. 34, Fig. 35) present the nonverbal behavior displayed by above categorized High, Low and Middle (where applicable) Choice head + body stimuli of Experiment 3.

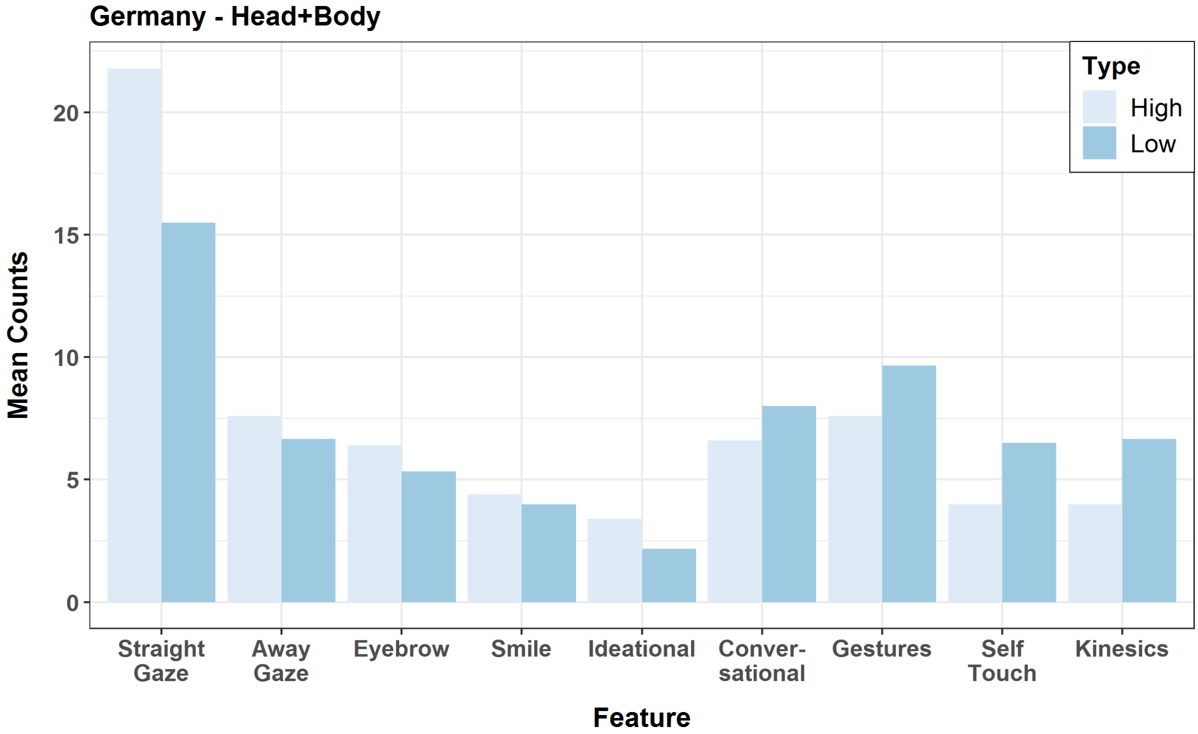

Fig. 34: Head+Body. Mean counts of gazing straight, gazing away, eyebrow movements, smiling, ideational gestures, conversational gestures, ideational+conversational (Gestures), self-touch and kinesics for German experiment and high and low categorized stimuli.

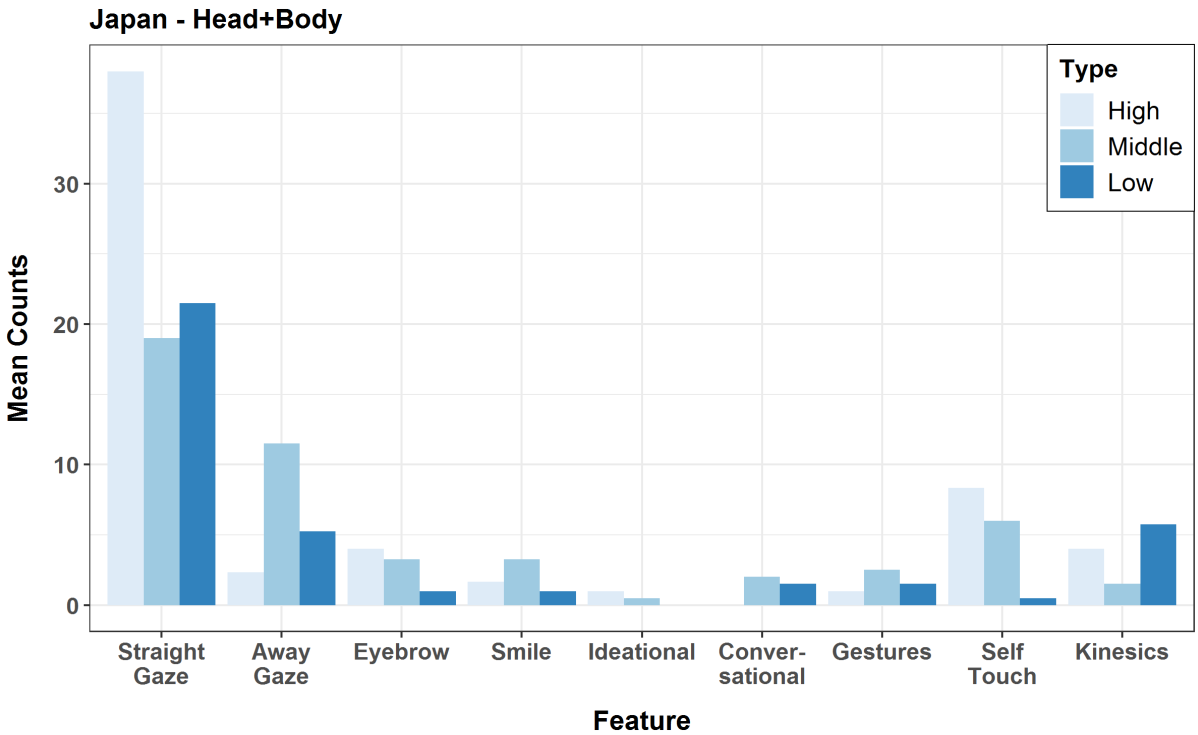

Fig. 35: Head+Body. Mean counts of gazing straight, gazing away, eyebrow movements, smiling, ideational gestures, conversational gestures, ideational+conversational (Gestures), self-touch and kinesics for Japanese experiment and high middle and low categorized stimuli.

Exploration of nonverbal behaviour of head+body stimuli revealed that in comparison to the low group preferred German interaction partner tended to display (Fig. 34):

- more straight gaze (high = 21.80, low = 15.50)
- approximately equal amounts of averted gaze (high = 7.60, low = 6.67)
- slightly more eyebrow movements (high = 6.40, low = 5.33)
- slightly more smiling (high = 4.40, low = 4.0)
- slightly less gestures (high = 7.60, low = 9.67) and self-touch (high = 4.0, low = 6.50)
- slightly more ideational gestures (high = 3.40, low = 2.17)
- slightly less conversational gestures (high = 6.60, low = 8.0)
- slightly less kinesics (high = 4.0, low = 6.67)

Japanese preferred interaction partner on the other hand tended to show (Fig. 35):

- more straight gaze (high = 38.0, middle = 19.0, low = 21.50)
- less averted gaze (high = 2.3, middle = 11.50, low = 5.25)
- slightly more eyebrow movements (high = 4.0, middle = 3.25, low = 1.0)
- no obvious trend for smiling (high = 1.67, middle = 3.25, low = 1.0)
- little gestures in general (high = 1.0, middle = 2.5, low = 1.5)
- little ideational (high = 1.0, middle = 0.5, low = 0.0)
- little conversational gestures (high = 0.0, middle = 2.0, low = 1.5)
- more self-touch (high = 8.33, middle = 6.0, low = 0.50)
- moderate amount of kinesics (high = 4.0, middle = 1.50, low = 5.75)

For the sake of completeness also nonverbal behavior displayed by the separate head + body stimuli of Experiment 3 ordered by choice according to Dunn's test is presented (see Fig. 36).

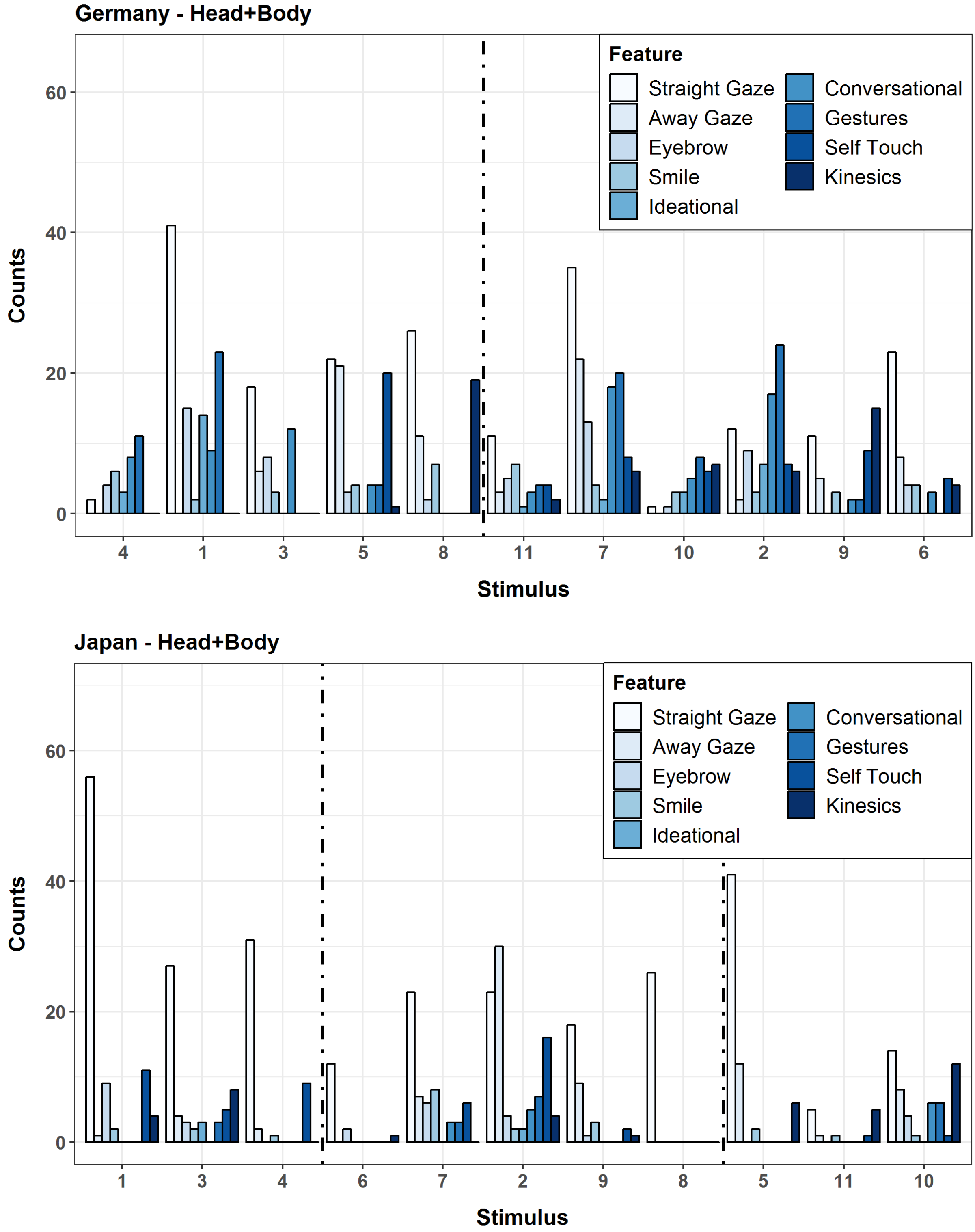

Fig. 36: Head+Body. Counts of gazing straight and away, eyebrow movements, smiling, ideational gestures, conversational gestures, ideational+conversational (Gestures), self-touch and kinesics for every stimulus for German (upper) and Japanese (lower) experiment. Dashed line indicates the boundary between high(left), middle (middle, where applicable) and low (right) categorized stimuli.

4.4 Discussion

In this study I investigated in more detail the social judgment of potential interaction partner for future cooperation based on nonverbal cues. For that purpose, participants in Germany and Japan viewed video stimuli of potential cooperation partner showing only their face, only their body and both face and body. The goal for these 3 experiments was to answer the following three research questions:

First, I aimed at examining the importance of facial expressions by comparing judgments from the first experiment, head stimuli, and the third experiment, both head and body stimuli. I hypothesized that although information from the head is important, body language might play a much higher role than the focus on face perception in the scientific community might imply. Thus, the perception of facial expression may not necessarily be the pivotal element for making social decisions on interaction partner’s trustworthiness.

Secondly, I focused on cultural differences of body language in German and Japanese culture. As it has been shown that humans seem to be biased in favor of the faces of their own culture (Ratner et al., 2014). I wanted to find out if a similar tendency would be observed for body language. Hence, I compared if in German and Japanese culture participants would preferably choose people from their own culture based on watching only their body language.

Lastly, I investigated whether Japanese and German participants tended to preferably choose one specific person or group of people as future interaction partner and thus perceived them as more trustworthy. I explored what nonverbal behavior these chosen people showed and whether there existed common behaviors among them also in comparison to the people that weren’t chosen.

4.4.1 The Importance of Facial Expression

The comparison of Experiment 1, head stimuli, and Experiment 3, whole body, showed that only a minority of participants, that is only about a third of them, choose the same person based on watching only the head as compared to watching the whole body (with head) (see Fig. 24). Importantly, this result remained the same when including participants who also choose the same person in Experiment 2, based on solely the body, as in Experiment 1, based on solely the head. Thus, it becomes clear that although facial expression is important for making social judgments other nonverbal bodily cues, such as body kinesics or gestures, may at least be equally as important. In fact, the combination of face and body movements may be a deciding factor here.
In this light, studies on physiognomics and facial expression perception alone may not be of such high relevance in real life, where we are always confronted with the whole body of the interaction partner. For future research, it would be interesting to test if humans have an equally fine-tuned

perception mechanism for body movements as has been described by Haxby and colleagues (2000, 2002) for the face. Indeed various studies showed, by using point-light displays, that people may be able to distinguish emotional expressivity and make social judgments based on body actions, which implies that humans' perception seems to be sensitive enough to characteristics of body movements (Atkinson et al., 2004; Chouchourelou et al., 2006; Dael et al., 2012; De Gelder et al., 2015; Montepare & Zebrowitz-McArthur, 1988).

4.4.2 Cultural Bias for Body Language

Interestingly, results showed that there may be cultural differences between Japan and Germany in social judgment based on in-group, meaning own culture's, versus outgroup, meaning other culture's body language (Fig. 25). German participants seemed to prefer body language from their own culture slightly more than Japanese participants. Nevertheless, as expected, also Japanese participants preferably choose interaction partner based on body movements of their own culture as compared to other culture. Thus, one may conclude that in general own culture's body language is preferred, presumably because it is more familiar. This would be in line with similar research on face perception, where it has been continuously shown that own culture bias is prevalent in memory tasks (see review: Meissner & Brigham, 2001) as well as social judgment tasks (Ratner et al., 2014) and that the origin of this bias may lie in early childhood experience (Scott & Monesson, 2009).

4.4.3 Nonverbal Behavior of Perceived Trustworthiness

After comparing the choice scores of all head stimuli, body stimuli and head + body stimuli, results showed that participants from both Japan and Germany tended to prefer one person, or a group of persons, as their interaction partner (Fig. 26 - Fig. 29). In Japan in particular, results for head as well as body stimuli showed one or two clear "favorites" and even in the whole-body (head+body) experiment the preferred group tended to be small. In comparison, German participants' preferred group seemed to be larger and less well specified in all three categories, head, body, head + body. When exploring in more detail the exact nonverbal behaviors of persons that seemed to be regarded as more trustworthy, without statistical analysis owing to the small group sizes, a few common denominators seemed to emerge, particularly regarding the features gaze behavior, gestures and kinesics.

For German participants, when judgment of interaction partner occurred based on the head alone, straight gaze didn't seem to be a critical feature, as even less preferred stimuli tended to look straight and overall even more so than the preferred group. Similarly, averted gaze, eyebrow movements and smiles tended to be equally displayed by preferred as well as less preferred stimuli. Overall it appeared that moderate amounts of straight and averted gaze and few eyebrow movements and smiles tended to be more highly regarded for German interaction partner (Fig. 30,

Fig. 31). However, in Japan head stimuli that showed high amount of straight gaze and little to none averted gaze tended to be preferred, whereas smile and eyebrow movements might not play a big role (Fig. 30, Fig. 31). Most interestingly, for Japanese participants this pattern remained even when the head+body stimuli were displayed (Fig. 35), possibly owing to the fact that Japanese participants in general focused more on the face even when the body information was available, similar to evidence shown in Chapter 3. Furthermore, when looking at the preferred interaction partner separately it seems not much changed when body information became available (Fig. 31, Fig. 36). In detail, Japanese participants tended to judge interaction partner 1,3 and 4 highly and 6 moderately high when viewing only the head and in the same manner they judged interaction partner 1,3 and 6 highly and 4 moderately high when viewing the head and the body (Fig. 31, Fig. 36). In contrast, for German participants preferred interaction partner changed when both head and body information was available, in comparison to the situation when judgment was based solely on the head (Fig. 31, Fig. 36). Thus, it becomes clear that body information may have played a bigger influence on the decision for the German cooperation partner as compared to the Japanese cooperation partner, where information from the head, possibly the gaze was most informative.

In terms of gestures, in Germany there seemed to emerge a trend of a preference for interaction partner with less self-touch (Fig. 32, Fig. 34), which is in line with previous research that showed presenters performing self-touch were less effective in terms of communication style (Maricchiolo et al., 2009). In Japan on the other hand the opposite trend appeared with preferred partner also displaying self-touch (Fig. 32, Fig. 35). This might again be evidence in favour of the hypothesis that Japanese people focus less on body movements than on facial expressions, thus even showing self-touch wouldn't influence judgment. However, this trend needs to be further tested in future research.

Also, although in Germany a preference for highly gesticulating interaction partner may have become apparent from results on body stimuli (Fig. 32), this preference wasn't repeated, when both, head and body information, was available (Fig. 34). Instead, the opposite trend arose, that is preferred people showing slightly less gestures. However, ideational gestures still tended to be presented slightly more by the preferred group rather than the less preferred group. This is similar to results shown by Maricchiolo and colleagues (2009) who demonstrated in a task on message persuasiveness that ideational gestures were in general most positively regarded.
In Japan on the other hand gestures in general tended not to be displayed much by the Japanese experts as becomes apparent from results of Experiment 3 – Head + Body (Fig. 36) and even in Experiment 2 where German and Japanese stimuli were intermixed, highly gesticulation German stimuli were found in the highly preferred (G1, G7) and the less preferred group (G2) (Fig. 33). Thus,

this demonstrates that gesture rate overall may not be a critical factor for Japanese decision makers. Kinesics, on the other hand tended to be important for Japanese decision makers, but again only when body stimuli alone were presented, as can be seen from the trend of low judgment of interaction partner with a high kinesics count. One could perhaps describe them as the more fidgeting interaction partner (Fig. 32). Similarly, although less high in general the kinesics score tended to be also a little higher for the less preferred group when viewing combined head and body stimuli (Fig. 35) and particularly the person in last place (10) showed the highest kinesics score (Fig. 36).

Of course, it would be relevant for future research to consider implementing findings from this explorative analysis in a study with embodied experts. The same is true also for results of this study's research question 1, showing the influence of body motion in judgment of trustworthiness, and research question 2, on the in-culture body motion bias. This is in line with hypotheses on second-person neuroscience about the importance of being emotionally engaged and interacting with a person to simulate a natural social situation where specific social cognitive resources are recruited (Schilbach et al. 2013). However, it has also been demonstrated that in some instances even screen-based stimuli may be able to elicit activation of cognitive brain areas typical for social interaction. Schilbach and colleagues (2006) showed that using self-directed avatars on a screen displaying nonverbal behavior as opposed to other-directed avatars seemed to elicit activation of "social brain" areas that have been implicated in emotional processing. Thus, being looked at directly by screen-based human-like stimuli, in contrast to observing someone from a third-person perspective, may already elicit some emotional engagement and in turn already simulate at least in part the necessary conditions of a naturalistic interaction. In the same vein, one may argue that also in this study, where the experts shown addressed the participants directly certain social cognitive resources were recruited and at least in part a natural social judgment occurred. This may have been further facilitated through the expectation of participants to be rewarded when making a good choice, that is when choosing an expert that can be considered trustworthy. Hence, participants may have been motivated to understand the experts' minds based on their behavior, which might have led to a certain degree of emotional engagement, as according to second-person neuroscience (Schilbach et al., 2013) this may be a prerequisite for understanding others in social interaction.

4.4.4 Summary and Conclusion

In conclusion, with three experiments I showed that body movements were highly important when judging trustworthiness and choosing a partner for interaction. Similarly, I showed that there might exist a bias towards own versus other culture regarding body language. Furthermore, I demonstrated that based on their nonverbal behavior some persons seem to be more highly trusted and thus chosen for cooperation than others and I explored specific trustworthiness signals that may be essential when choosing an interaction partner. These signals tended to differ not only in terms of chosen and not chosen interaction partner, but also in terms of culture. Some evidence even indicates a trend that gaze behavior may be more important for judgments in Japanese culture and body signals may be more important in German culture. Importantly, this would be in line with results addressed in Chapter 3 that showed attention of Japanese participants to be focused more on the face rather than the body of interaction partner as compared to German participants' attention, which was focused on body as well as the face. However, this effect needs to be tested in future research.

Chapter 5 Robotics Outlook with Conclusion

5.1 Robots Enter Social Environment

So far in this work, I have detailed mechanisms and behaviors related to human social interaction. However, at present and presumably also in the near future a new player has entered the social stage. It is the machine. Due in part to technological advances and also bearing in mind the rising body of science fiction literature interest in the topic of "cross-species", that is human-machine or more specific human-robot interaction has been increasing. Interestingly, even before the advent of robotics humans tended to treat computers as social entities to which human-like norms apply (Nass & Moon, 2000; Nass et al., 1994). Similarly, anthropomorphism, in other words applying human-like characteristics to inanimate objects, has been a well-known concept for many years. It has been suggested that this type of behavior is based on the strategy of perceiving the world in terms of what is most important to oneself, which for humans oftentimes mean humans (Guthrie, 1997). Likewise, the intentional stance theory (Dennett, 1989), a strategy wherein humans treat systems as intentional agents may be a plausible explanation for anthropomorphism (Hayes & Miller, 2010).

5.2 Benefit of Nonverbal Behavior for Human-Robot Interaction (HRI)

Thus, anthropomorphism has been considered useful for robots to better integrate into the social sphere (Duffy, 2003). Nowadays, with more sophisticated technology research in human-robot interaction (HRI), concerning appearance and behavior of the robot has progressed. Particularly, in contexts where robots come into close contact with humans, adopting human-like social behavior seems to be of advantage for smooth social interaction (Dautenhahn, 2007). Many research groups have applied this idea by incorporating human facial expressions (Breazeal, 2003), body movements (Takayama, Dooley, & Ju, 2011) gestures (Ono, Imai, & Ishiguro, 2001) and gaze cues (Breitfuss, Prendinger, & Ishizuka, 2008; Kompatsiari, Tikhanoff, Ciardo, Metta, & Wykowska, 2017) into robots' behavior during HRI. Emotional expressivity displayed by the robot in particular has been shown to improve enjoyment in the interaction for adults (Bartneck, 2003) and children (Tielman, Neerincx, Meyer, & Looije, 2014). Additionally, cooperation and performance in a team-work related task has been shown to increase when a robot employed nonverbal cues, such as head, gaze and body movements (Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005). Similarly, children in a robot-tutoring task interacting with a robot, who showed increasingly socially supportive nonverbal behavior, improved their test scores in comparison to an interaction with a neutral robot (Saerbeck, Schut, Bartneck, & Janse, 2010). In this vein, attempts have been made to adapt the personality of a robot, by modulating gesture rate and head pose, to the human partner with evidence indicating a preference of participants towards a robot that matched their personality (Aly & Tapus, 2013). All these examples show what has long been suggested, namely that human-robot interaction and cooperation may benefit from implementation of human nonverbal cues. Thus, robots may well be able to not only imitate humans, but also use as sophisticated a social mechanism as joint attention, a building block of cooperation (Admoni & Scassellati, 2015; Bütepage & Kragic, 2017).

5.3 Trust in Robots

Trust in robots is another factor that has been discussed in depth to influence successful human-robot interaction and cooperation. Although, in this domain performance has consistently shown to be one of the most important determinants, robot attributes also play an essential role in perceived trustworthiness (Cameron et al., 2015; Hancock et al., 2011; Lewis, Sycara, & Walker, 2018). In this sense various studies have examined the impact of robots behavior, such as gaze, body (van den Brule, Dotsch, Bijlstra, Wigboldus, & Haselager, 2014) and head movements (Złotowski et al., 2016) as well as gestures (Van den Brule, Bijlstra, Dotsch, Wigboldus, & Haselager, 2013) on human trust. However, many studies didn't find a robust impact of nonverbal behaviors on perceived trustworthiness (Stanton & Stevens, 2017; Van den Brule et al., 2014). Nevertheless, negative headshake as feedback in combination with machine-like as compared to human-like appearance of a robot seemed to decrease attribution of trustworthiness (Złotowski et al., 2016). Furthermore, gaze contact during a difficult decision task seemed to increase trustworthiness in the robot in comparison to gaze contact during an easy decision task (Stanton & Stevens, 2014).

5.4 Culture and Robots

Culture is an additional aspect which has been considered to modulate perceived trustworthiness in automation as well as artificial agents (Du, Tilbury, Robert, Yang, & Pradhan, 2018; Huang & Bashir, 2017; Lewis et al., 2018; Yerdon, Marlowe, Volante, Li, & Hancock, 2017). However, not only trustworthiness but also acceptance of the robot seems to be influenced by culture (Lee & Sabanović, 2014; Šabanović, 2010; Šabanović, Bennett, & Lee, 2014). Therefore, nonverbal behavior may also play a role, as it has been shown that matching gestures of an artificial agent to different cultures increased acceptability (Hayes & Miller, 2010; Maldonado & Hayes-Roth, 2004). In the study of cultural robotics Japan in particular has been of interest, as Japanese culture has long been seen as a robot-loving culture, although clever application of marketing and image tending may have played a role here (Šabanović et al., 2014). Nevertheless, it is obvious that humanoid robots for consumers, particularly with the launch of robot Pepper are becoming more common in Japan.

5.5 Conclusion and Outlook for Robotics

As human social norms have been regarded as a template for robot behavior and because robots are built to share the human environment, research into human interaction may be used to also inform human-robot interaction. Thus, the results reported in this work may be applicable also for robotics. This seems even more important when considering that "good", i.e. correctly applied, nonverbal behavior may improve HRI, and "wrong" nonverbal behavior can also negatively affect interaction (see review Admoni & Scassellati, 2015). In this way, it becomes clear that studying effects of human-human interaction before making robots behave in an arbitrarily "human way" is essential.

Chapter 1 discussed how humans respond to irrelevant pointing gestures and deictic gaze during real life interaction. Previously it has been suggested these directional social cues would elicit automatic gaze following (Friesen & Kingstone, 1998). Indeed, directional gaze and pointing movements have also been applied to social robotics experiments (see reviews Admoni & Scassellati, 2015, 2017). However, I propose that one needs to be careful of the context when applying these nonverbal signals in an ecologically valid real life HRI scenario. My results showed that directional cues may not capture attention if there is no meaningful pointing intention, for example a specific target, conveyed by the social signal. The evidence in the experiments reported here suggests that observers remain fixated on the face of the interaction partner (Fig. 5, Fig. 9).

Based on results presented in this work, one can conclude that in similar HRI scenarios robots should, on the one hand, follow directional gestures and gaze only in contexts where the interaction partner has an explicit target in mind, but should remain fixated on the humans face in other cases, for example where pointing and gaze shift seem to have regulatory function in terms of conversation flow. Similarly, a robot making gestures and gaze shifts may only expect these to elicit attentional capture by the human observer if there is a target or a task relevant to the pointing.

In general, my study also showed the importance of the face as a focus of attention during interaction. Hence, it may not be amiss to incorporate a face into the design of a robot partner, with the expectation that a human would try to read many social cues from it. Also, when designing a robot with new or unusual nonverbal behavior that a human partner needs to learn to decipher, putting it into the head area of the robot would probably be easiest for the human to decode rather than imposing an absolutely new habit of focusing attention elsewhere.

Lastly, I discussed the dichotomy of my results not eliciting attentional capture in comparison to other paradigms studying similar social but non-embodied stimuli in the laboratory (Friesen & Kingstone, 1998) and came to the conclusion that in real life, signals are never seen in isolation. Therefore, when considering the abundance of provided signals a human may send, one specific directional cue may have on the whole less of an impact on the observer. Similarly, one should consider that a robot during HRI may need to send multiple nonverbal signals, otherwise in the absence of all others, one signal might receive more weight and immediately capture attention unnecessarily.

Chapter 2 examined how people from different cultures, specifically in Germany and Japan, react in terms of gaze behavior while listening, being asked a question and answering. Results showed that while listening and being asked a question no difference in gaze behavior between the two cultures occurred (Fig. 14). Instead, participants mostly seemed to remain fixated on the face, which was similar to the culturally robust behavior elicited by directional cues in Chapter 1. However, while answering Japanese participants seemed to be equally likely to avert gaze than to look at their interaction partner, which contrasted with behavior of the German sample, who mostly looked the interaction partner in the face, also while speaking (Fig. 15). Social norms of German participants to keep eye contact rather than look away while thinking or also Japanese need for decoding of facial expression falling away during speaking may be a few of the reasons for these differences.
It might thus be beneficial for a robot to adapt in a similar way so that in Germany it keeps looking at the face also when speaking and thinking about an answer or averts gaze in Japan when answering a question.
Also, these results may be further evidence that in Japan when listening to a conversation partner, nonverbal cues, particularly of the face may be highly important for understanding. This may in turn be important for purely voice based artificial agents, such as Siri and all its future "siblings", as in Japan embodiment may be necessary for communication.

Of note for robotics also the results on nonverbal signals of trustworthiness and decisions about cooperation partner, as described in Chapters 3 and 4, might be relevant. Especially, when keeping in mind the importance of trust in any interaction (see Introduction, Chapter 3 – Introduction), but also in HRI (see Chapter 5 – Trust in robots).
Chapter 3 investigated gaze behavior and attention towards trustworthy interaction partner, but also how culture shaped this gaze behavior. Although participants seemed to show the same gaze patterns when confronted with a person they trusted and choose to cooperate with rather than a person they trusted less and would not want to cooperate with, results did show a cultural difference (Fig. 19). As expected from results of Chapters 1 and 2 on listening behavior, also when judging presenters on trustworthiness, Japanese participants looked mostly at the face.
However, in contrast to results of Chapter 2, German participants looked less at the face. Thus, it seems that when judging trustworthiness and cooperation and particularly when the verbal channel is not available or non-informative German participants may also take other information into account. On the other hand, in Japanese culture the body did not seem to be informative enough to

base their decisions on.
Hence one may conclude that in Germany other information channels than the face, for example gestures and body motion, should be included in robot design, whereas for Japan the face area, that is facial expression and gaze seem to be much more important.

Understanding the specific type of body and face movement has been the aim of Chapter 4. I showed that participants of both cultures, but more so in Germany, seem to have a positive bias towards body movements including gestures and kinesics from their own culture (Fig. 25). This again demonstrates the importance of nonverbal behavior incorporated into robotics being adjusted to culture.

Importantly, results of Chapter 4 also indicated that gestures and body motion in general seem to be essential when choosing a cooperation partner, as participants choose a different cooperation partner in approximately two thirds of the time when information about the whole body was available than when only the face was visible (Fig. 24).
In part, this is consistent with findings from real human interaction (Chapter 3) that demonstrated how in Germany the face area may be of slightly less interest as a nonverbal signal of trustworthiness as compared to Japan.
Interestingly, contrary to expectations from Chapter 3 this behavior also occurred in Japanese participants, that is Japanese participants similarly changed their choice of interaction partner when getting information about the whole body. Thus, it seems that although in real interactions Japanese participants do not make many overt gaze movements towards other body areas, body motion may still be something to be considered. However, in the study explained in Chapter 3 only four stimuli were presented for judgment to Japanese raters and no one was chosen more than the others. Thus, it may be possible that Japanese participants focus their attention mostly on the face but do shift attention only in cases when unusual or untrustworthy behavior is detected.
Furthermore, when looking at specific nonverbal signals that may be displayed by chosen interaction partner in Japan, gaze behavior, that is straight rather than averted gaze, and possibly also standing securely without fidgeting tend to be the critical points for decision (Fig. 30,Fig. 32, Fig. 35). Gestures however don’t tend to be a decisive factor in Japan, whereas in Germany the degree of self-touch and possibly also the amount of ideational gestures (Fig. 32, Fig.34) might have influenced decision, which is similar to previous findings reported by Maricchiolo et al. (2009) on persuasiveness. Gaze behavior or other facial expressions on the other hand did not tend to differ much between preferred and less preferred cooperation partner in Germany (Fig. 34).
Concerning robotics, it becomes apparent that especially for Japan, gaze contact during initial introduction and first impression formation tends to be highly regarded and therefore should

perhaps be implemented in a social robot to be considered trustworthy. However, in Germany a mixture of moderate straight and averted gaze may be of equal advantage and particularly some ideational gestures may be useful.

In summary, nonverbal behavior, particularly gaze seem to be dependent not only on context, but also on culture. However, some basic cognitive mechanisms such as attentional capture which triggers gaze following appear to be culturally robust. In this case it is of note that embodied interaction may differ from merely observing 2D stimuli, which should be kept in mind for future research into gaze behavior and their underlying cognitive mechanisms, as real life is always a mixture of various signals. In the same vein, gestures and body motion can be assumed to be important for attribution of trustworthiness and shouldn’t fall behind the facial expression research. Interestingly, judgments on trustworthiness and the nonverbal behaviors involved may be dependent on culture. In Japan, the face and the gaze may be more crucial, whereas in Germany the body may be more informative. Lastly, as nonverbal behavior is an essential part of HRI, it is important to implement correct behavior into robots to elicit the desired response of a smooth interaction with a human. Thus, applying knowledge gained from human-human interaction into HRI may be of value. Furthermore, as it has been shown in this work certain culturally specific behaviors should also be kept in mind when designing a social robot.

References

Admoni, H., & Scassellati, B. (2015). *Robot Nonverbal Communication as an AI Problem (and Solution).* Paper presented at the AAAI 2015 Fall Symposium Series, Arlington, VA.

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *Journal of Human-Robot Interaction, 6*(1), 25-63.

Adolphs, R. (2002). Trust in the brain. *Nature Neuroscience, 5*(3), 192.

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences, 1191*(1), 42-61.

Aly, A., & Tapus, A. (2013). *A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction.* Paper presented at the 8th ACM/IEEE international conference on Human-robot interaction. Piscataway, NJ: IEEE Press.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*(3), 431.

Ando, S. (2004). Perception of gaze direction based on luminance ratio. *Perception, 33*(10), 1173-1184.

Argyle, M. (1988). *Bodily communication* (2nd ed.). London, England: Routledge

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

Argyle, M., & Kendon, A. (1967). The Experimental Analysis of Social Performance. In L.Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 3, pp. 55-98). New York, NY: Academic Press.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*(5), 485-496.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and the theory of mind*. Cambridge, MA: MIT Press/Bradford Books.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a" language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual cognition, 4*(3), 311-331.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord, 31*(1), 5-17.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry, 42*(2), 241-251.

Bartneck, C. (2003). *Interacting with an embodied emotional character.* Paper presented at the 2003 international conference on Designing pleasurable products and interfaces. New York, NY: ACM.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication, 52*(3), 566-580.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition: International Journal of Cognitive Science, 104*(3), 644-653.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science, 17*(6), 514-520.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior, 10*(1), 122-142.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Get real! Resolving the debate about equivalent social stimuli. Visual cognition, 17(6-7), 904-924.

Bond, M. H., & Komai, H. (1976). Targets of gazing and eye contact during interviews: Effect on Japanese nonverbal behavior. *Journal of Personality and Social Psychology, 34*(6), 1276.

Bonoma, T. V. (1976). Conflict, Cooperation and Trust in 3 Power-Systems. *Behavioral Science, 21*(6), 499-514.

Boone, R. T., & Buck, R. (2003). Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *Journal of Nonverbal Behavior, 27*(3), 163-182.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International journal of human-computer studies, 59*(1-2), 119-155.

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.* Paper presented at the 2005 IEEE/RSJ international conference on intelligent robots and systems, Edmonton, Canada.

Breitfuss, W., Prendinger, H., & Ishizuka, M. (2008). Automatic generation of gaze and gestures for dialogues between embodied conversational agents. *International Journal of Semantic Computing, 2*(01), 71-90.

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: a review. *Gerontechnology, 8*(2), 94-103.

Bütepage, J., & Kragic, D. (2017). Human-robot collaboration: From psychology to social robotics. *arXiv preprint arXiv:1705.10146*.

Cabibihan, J.-J., Javed, H., Ang, M., Jr., & Aljunied, S. (2013). Why Robots? A Survey on the Roles and Benefits of Social Robots in the Therapy of Children with Autism. *International Journal of Social Robotics, 5*(4), 593-618.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527.

Cameron, D., Aitken, J., Collins, E., Boorman, L., Chua, A., Fernando, S., . . . Law, J. (2015). *Framing factors: the importance of context and the individual in understanding trust in human-robot interaction.* Paper presented at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2015. Piscataway, NJ: IEEE Press.

Cassell, J., & Bickmore, T. (2000). External manifestations of trustworthiness in the interface. *Communications of the ACM, 43*(12), 50-56.

Chouchourelou, A., Matsuka, T., Harber, K., & Shiffrar, M. (2006). The visual analysis of emotional actions. *Social Neuroscience, 1*(1), 63-74.

Cole, G. G., Skarratt, P. A., & Kuhn, G. (2016). Real person interaction in visual attention research. *European Psychologist*, *21*, 141–149.

Cook, M. (1977). Gaze and Mutual Gaze in Social Encounters: How long—and when—we look others "in the eye" is one of the main signals in nonverbal communication. *American Scientist, 65*(3), 328-333.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12*(5), 1085.

Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362*(1480), 679-704.

De Gelder, B., De Borst, A., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 149-158.

Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA, USA: MIT press.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111*(2), 203.

DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science, 23*(12), 1549-1556.

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition, 33*(4), 727-733.

Dovidio, J. F., & Ellyson, S. L. (1982). Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly, 45*(2), 106-113.

Dovidio, J. F., & Ellyson, S. L. (1985). Pattern of Visual Dominance Behavior in Humans. In S. L. Ellyson & J. F. Dovidio (Eds.), *Power, Dominance, and Nonverbal Behavior* (pp. 129-149). New York, NY: Springer New York.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual cognition, 6*(5), 509-540.

Du, N., Tilbury, D., Robert, L., Yang, X. J., & Pradhan, A. (2018). *A Cross-Cultural Study of Trust Building in Autonomous Vehicles.* Paper presented at the Conference on Autonomous Vehicles in Society: Building a Research Agenda, East Lansing, MI.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems, 42*(3-4), 177-190.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology, 58*(2), 342.

Ekman, P., & Friesen, W. V. (1969a). Nonverbal leakage and clues to deception. *Psychiatry: Interpersonal and Biological Processes, 32*(1), 88-106.

Ekman, P., & Friesen, W. V. (1969b). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica: Journal of the International Association for Semiotic Studies/Revue de l'Association Internationale de Sémiotique, 1*(1), 49-98.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124.

Ekman, P., Friesen, W. V., & Tomkins, S. S. (1971). Facial affect scoring technique: A first validity study. *Semiotica: Journal of the International Association for Semiotic Studies/Revue de l'Association Internationale de Sémiotique, 3*(1), 37-58.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*(3875), 86-88.

Ellsworth, P. C., & Ludwig, L. M. (1972). Visual behavior in social interaction. *Journal of Communication, 22*(4), 375-403.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews, 24*(6), 581-604.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19*(9), 1508-1519.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99*(14), 9602-9605.

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience, 16*(8), 1320-1326.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. Vision Research, 51(17), 1920-1931.

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6*(3), 534.

Freeman, J. B., Rule, N. O., Adams Jr, R. B., & Ambady, N. (2009). Culture shapes a mesolimbic response to signals of dominance and subordination that associates with behavior. *NeuroImage, 47*(1), 353-359.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*(3), 490-495.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*(4), 694.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *The American Journal of Psychology*, *76*, 86–94.

Gobel, M. S., Kim, H. S., & Richardson, D. C. (2015). The dual function of social gaze. *Cognition: International Journal of Cognitive Science, 136*, 359-364.

Gordon, D. S., & Platek, S. M. (2009). Trustworthy? The brain knows: Implicit neural responses to faces that vary in dark triad personality characteristics and trustworthiness. *Journal of Social, Evolutionary, and Cultural Psychology, 3*(3), 182.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality, 37*(6), 504-528.

Guthrie, S. E. (1997). Anthropomorphism: A definition and a theory. In *Anthropomorphism, anecdotes, and animals.* (pp. 50-58). Albany, NY: State University of New York Press.

Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text & Talk, 26*(3), 281-328.

Hall, E. (1976). *Beyond Culture*. New York, NY: Anchor Press.

Hall, E. T., & Hall, M. R. (2001). Key concepts: Underlying structures of culture. *International HRM: Managing diversity in the workplace*, 24-40.

Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. *Psychological Bulletin, 131*(6), 898.

Hamilton, S. J. (2017). The Effects of Pointing Gestures on Visual Attention. *University Honors Program Theses, 243*.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517-527.

Hardin, R. (2002). *Trust and trustworthiness*. New York, NY, USA: Russell Sage Foundation.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*(6), 223-233.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological psychiatry, 51*(1), 59-67.

Hayes, C. C., & Miller, C. A. (2010). *Human-computer etiquette: Cultural expectations and the design implications they place on computers and technology*. Boca Raton, FL: CRC Press.

Hayward, D. A., Voorhies, W., Morris, J. L., Capozzi, F., & Ristic, J. (2017). Staring reality in the face: A comparison of social attention across laboratory and real world measures suggests little common ground. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 71*(3), 212.

Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: The effects of gaze direction on face recognition in children and adults. *Developmental science, 6*(1), 67-71.

Huang, H.-Y., & Bashir, M. (2017). *Users' Trust in Automation: A Cultural Perspective.* Paper presented at the International Conference on Applied Human Factors and Ergonomics.

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: the core of social cognition. *Neuroscience & Biobehavioral Reviews, 33*(6), 843-863.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General, 141*(1), 19.

Jarick, M., & Kingstone, A. (2015). The duality of gaze: eyes extract and signal social information during sustained cooperative and competitive dyadic gaze. *Frontiers in Psychology, 6*(1423).

Kajopoulos, J. (2015). *A mobile eye-tracking study: Cultural differences in gaze patterns during social interaction.* (Master thesis), Ludwig-Maximilians Universität München, Munich, Germany.

Kajopoulos, J., Wong, A. H. Y., Yuen, A. W. C., Dung, T. A., Kee, T. Y., & Wykowska, A. (2015). Robot-Assisted Training of Joint Attention Skills in Children Diagnosed with Autism. In A. Tapus, E. André, J.-C. Martin, F. Ferland, & M. Ammi (Eds.), *Social Robotics.* Paper presented at ICSR 2015, Social Robotics: 7th International Conference (pp. 296-305). Cham, Switzerland: Springer International.

Katagiri, Y., Nass, C., & Takeuchi, Y. (2001). Cross-cultural studies of the computers are social actors paradigm: The case of reciprocity. *Usability evaluation and interface design: Cognitive engineering, intelligent agents, and virtual reality*, 1558-1562.

Kaul, T. J., & Schmidt, L. D. (1971). Dimensions of interviewer trustworthiness. *Journal of Counseling Psychology, 18*(6), 542.

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

Kidwell, M. (2009). Gaze shift as an interactional resource for very young children. *Discourse processes, 46*(2-3), 145-160.

Kinsbourne, M. (1972). Eye and Head Turning Indicates Cerebral Lateralization. *Science, 176*(4034), 539-541.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes, 24*(2), 145-167.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological Bulletin, 100*(1), 78.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*(5), 419-435.

Kompatsiari, K., Ciardo, F., Tikhanoff, V., Metta, G., & Wykowska, A. (2018). On the role of eye contact in gaze cueing. *Scientific Reports, 8*(1), 17842.

Kompatsiari, K., Tikhanoff, V., Ciardo, F., Metta, G., & Wykowska, A. (2017). *The Importance of Mutual Gaze in Human-Robot Interaction.* Paper presented at the International Conference on Social Robotics, Tsukuba, Japan. Cham, Switzerland: Springer International.

Koppensteiner, M., Stephan, P., & Jäschke, J. P. M. (2015). More than words: Judgments of politicians and the role of different communication channels. *Journal of Research in personality, 58*, 21-30.

Koppensteiner, M., Stephan, P., & Jäschke, J. P. M. (2016). Moving speeches: Dominance, trustworthiness and competence in body motion. *Personality and Individual Differences, 94*, 101-106.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In *Advances in experimental social psychology* (Vol. 28, pp. 389-450). New York, NY: Academic Press.

Krumhuber, E., Manstead, A. S., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*(4), 730.

Kurzban, R. (2001). The social psychophysics of cooperation: Nonverbal communication in a public goods game. *Journal of Nonverbal Behavior, 25*(4), 241-259.

Langton, S. R., & Bruce, V. (2000). You must see the point: automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 747.

Lee, D. Y., McGill, M. E., & Uhlemann, M. R. (1988). Counsellor and Client Reliance on Verbal and Nonverbal Cues in Judging Competency, Trustworthiness, and Atractiveness. *Canadian Journal of Counselling and Psychotherapy/Revue canadienne de counseling et de psychothérapie, 22*(1).

Lee, H. R., & Sabanović, S. (2014). *Culturally variable preferences for robot design and use in South Korea, Turkey, and the United States.* Paper presented at the ACM/IEEE international conference on Human-robot interaction. New York, NY: ACM.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social forces, 63*(4), 967-985.

Lewis, M., Sycara, K., & Walker, P. (2018). The role of trust in human-robot interaction. In *Foundations of Trusted Autonomy* (pp. 135-159). Cham, Switzerland: Springer International.

Lewis, M. B., & Edmonds, A. J. (2003). Face detection: Mapping human performance. *Perception, 32*(8), 903-920.

Lount Jr, R. B. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *Journal of Personality and Social Psychology, 98*(3), 420-433.

Lucas, G., Stratou, G., Lieblich, S., & Gratch, J. (2016). *Trust me: multimodal signals of trustworthiness.* Paper presented at the 18th ACM International Conference on Multimodal Interaction. New York, NY: ACM.

Macdonald, R. G., & Tatler, B. W. (2018). Gaze in a real-world social interaction: A dual eye-tracking study. Quarterly Journal of Experimental Psychology, 71(10), 2162-2173.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science, 13*(5), 460-464.

Maldonado, H., & Hayes-Roth, B. (2004). Toward cross-cultural believability in character design. *Agent Culture: Human-Agent Interaction in a Multicultural World*, 143-175.

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes, 24*(2), 239-266.

Mason, M., Hood, B., & Macrae, C. N. (2004). Look into my eyes: Gaze direction and person memory. *Memory, 12*(5), 637-643.

Mason, M. F., Tatkow, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological Science, 16*(3), 236-239.

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*(2), 143-157.

Maynard, S. K. (1987). Interactional functions of a nonverbal sign Head movement in japanese dyadic casual conversation. *Journal of Pragmatics, 11*(5), 589-606.

Maynard, S. K. (1990). Conversation management in contrast: Listener response in Japanese and American English. *Journal of Pragmatics, 14*(3), 397-412.

McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2006). Cultural Display rules drive eye gaze during thinking *Journal of cross-cultural psychology, 37*(6), 717-722.

McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2008). Gaze display when thinking depends on culture and context. *Journal of cross-cultural psychology, 39*(6), 716-729.

McDaniel, E. R. (1993). *Japanese Nonverbal Communication: A Review and Critique of Literature*. Paper presented at the International and Intercultural Division of the Speech Communication Association, Miami Beach, FL.

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *The American Journal of Psychology*, 311-323.

Mehrabian, A., & Williams, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *Journal of Personality and Social Psychology, 13*(1), 37.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3.

Montepare, J. M., & Zebrowitz-McArthur, L. (1988). Impressions of people created by age-related qualities of their gaits. *Journal of Personality and Social Psychology, 55*(4), 547.

Morey, R. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). Tutorials in Quantitative Methods for Psychology, 4.

Morsbach, H. (1973). Aspects of nonverbal communication in Japan. *Journal of Nervous and Mental Disease, 157*(4), 262-277.

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI-G). *European Journal of Psychological Assessment, 23*(3), 166-175.

Mulckhuyse, M., & Theeuwes, J. (2010). Unconscious attentional orienting to exogenous cues: A review of the literature. *Acta Psychologica, 134*(3), 299-309.

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception & Performance, 35*(1), 1-16.

Mundy, P. (2017). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*.

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

Nass, C., Steuer, J., & Tauber, E. R. (1994). *Computers are social actors.* Paper presented at the SIGCHI conference on Human factors in computing systems. New York, NY: ACM.

Nichols, K., & Champness, B. (1971). Eye gaze and the GSR. *Journal of Experimental Social Psychology, 7*(6), 623-626.

Ono, T., Imai, M., & Ishiguro, H. (2001). *A model of embodied communications with gestures between human and robots.* Paper presented at the 17th National Conference on Artificial Intelligence. Menlo Park, CA: AAAI.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087-11092.

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*(1), 128.

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190-206.

Pedersen, P. B. (2003). Culturally biased assumptions in counseling psychology. *The Counseling Psychologist, 31*(4), 396-403.

Perez-Osorio, J., Müller, H. J., Wiese, E., & Wykowska, A. (2015). Gaze following is modulated by expectations regarding others' action goals. *Plos one, 10*(11), e0143614.

Ratner, K. G., Dotsch, R., Wigboldus, D. H., van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality and Social Psychology, 106*(6), 897.

Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition: International Journal of Cognitive Science, 77*(1), B1-B14.

Robbins, B. G. (2016). What is Trust? A Multidisciplinary Review, Critique, and Synthesis. *Sociology Compass, 10*(10), 972-986.

Rossano, F. (2013). Gaze in conversation. In J. Sidnell & T.Stivers (Eds.), *The handbook of conversation analysis* (pp. 308-329). Chichester, England: Blackwell.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. *Conversation analysis: Comparative perspectives*, 187-249.

Šabanović, S. (2010). Robots in society, society in robots. *International Journal of Social Robotics, 2*(4), 439-450.

Šabanović, S., Bennett, C. C., & Lee, H. R. (2014). *Towards culturally robust robots: A critical social perspective on robotics and culture.* Paper presented at the ACM/IEEE HRI 2014 Workshop on Culture Aware Robotics (CARS), Bielefeld, Germany.

Sabatelli, R. M., & Rubin, M. (1986). Nonverbal expressiveness and physical attractiveness as mediators of interpersonal perceptions. *Journal of Nonverbal Behavior, 10*(2), 120-133.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). *Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor.* Paper presented at the SIGCHI conference on human factors in computing systems. New York, NY: ACM

Sato, W., Kubota, Y., Okada, T., Murai, T., Yoshikawa, S., & Sengoku, A. (2002). Seeing happy emotion in fearful and angry faces: qualitative analysis of facial expression recognition in a bilateral amygdala-damaged patient. *Cortex, 38*(5), 727-742.

Schaller, M. (2008). Evolutionary bases of first impressions. *First impressions*, 15-34.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and brain sciences, 36*(4), 393-414.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. Neuropsychologia, 44(5), 718-730.

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior, 31*(2), 87-94.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*(5), 402-409.

Scott, L. S., & Monesson, A. (2009). The origin of biases in face perception. *Psychological Science, 20*(6), 676-680.

Senju, A., Vernetti, A., Kikuchi, Y., Akechi, H., Hasegawa, T., & Johnson, M. H. (2013). Cultural background modulates how we look at other persons' gaze. *International Journal of Behavioral Development, 37*(2), 131-136.

Shepherd, S. V. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Frontiers in integrative neuroscience, 4*, 5.

Silk, J. B. (2002). Grunts, girneys, and good intentions: The origins of strategic commitment in nonhuman primates. In R. Nesse (Ed.), *Evolution and the capacity for commitment* (pp. 138-157). New York, NY: Russel Sage Press.

Silk, J. B., Kaldor, E., & Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour, 59*(2), 423-432.

Simmel, G. (1978). *The philosophy of money.* London, UK: Routledge & Kegan Paul, *236*, 232.

Sinha, P. (2000). Last but not least. *Perception, 29*(8), 1005-1008.

Skarratt, P. A., Cole, G. G., & Kingstone, A. (2010). Social inhibition of return. *Acta Psychologica, 134*(1), 48-54.

Skarratt, P. A., Cole, G. G., & Kuhn, G. (2012). Visual cognition during real social interaction. *Frontiers in Human Neuroscience*.

Stanton, C., & Stevens, C. J. (2014). *Robot pressure: the impact of robot eye gaze and lifelike bodily movements upon decision-making and trust.* Paper presented at the International Conference on Social Robotics. Cham, Switzerland: Springer International.

Stanton, C., & Stevens, C. J. (2017). Don't stare at me: the impact of a humanoid robot's gaze upon trust during a cooperative human–robot visual task. *International Journal of Social Robotics, 9*(5), 745-753.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*(3), 349-354.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., . . . Yoon, K.-E. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*(26), 10587-10592.

Symons, L. A., Hains, S. M., & Muir, D. W. (1998). Look at me: Five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development, 21*(3), 531-536.

Takayama, L., Dooley, D., & Ju, W. (2011). *Expressing thought: improving robot readability with animation principles.* Paper presented at the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI). New York, NY: ACM.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in cognitive sciences, 14*(8), 376-382.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 799-806.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65-70.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica, 135*(2), 77-99.

Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition: International Journal of Cognitive Science, 124*(3), 261-271.

Tielman, M., Neerincx, M., Meyer, J.-J., & Looije, R. (2014). *Adaptive emotional expression in robot-child interaction.* Paper presented at the 2014 ACM/IEEE international conference on Human-robot interaction. New York, NY: ACM.

Todorov, A. (2008). Evaluating faces on trustworthiness. *Annals of the New York Academy of Sciences, 1124*(1), 208-224.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social cognitive and affective neuroscience, 3*(2), 119-127.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813-833.

Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2013). Is cooperativeness readable in static facial features? An inter-cultural approach. *Evolution and Human Behavior, 34*(6), 427-432.

Töllner, T., Müller, H. J., & Zehetleitner, M. (2011). Top-down dimensional weight set determines the capture of visual attention: Evidence from the PCN component. *Cerebral Cortex, 22*(7), 1554-1563.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314-320.

Tsankova, E., Aubrey, A. J., Krumhuber, E., Möllering, G., Kappas, A., Marshall, D., & Rosin, P. L. (2012). *Facial and vocal cues in perceptions of trustworthiness.* Paper presented at the Asian Conference on Computer Vision. Berlin, Germany: Springer.

Uono, S., & Hietanen, J. K. (2015). Eye contact perception in the West and East: A cross-cultural study. *Plos one, 10*(2), e0118094.

Van den Brule, R., Bijlstra, G., Dotsch, R., Wigboldus, D. H., & Haselager, W. (2013). Signaling robot trustworthiness: Effects of behavioral cues as warnings. *Lecture Notes in Computer Science, 8239*, 583-584.

Van den Brule, R., Dotsch, R., Bijlstra, G., Wigboldus, D. H., & Haselager, P. (2014). Do robot performance and behavioral style affect human trust? *International Journal of Social Robotics, 6*(4), 519-531.

Van Mulken, S., André, E., & Müller, J. (1999). *An empirical study on the trustworthiness of life-like interface agents.* Paper presented at the 8th International Conference on Human-Computer Interaction (Vol. 2, pp. 152-156). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science, 28*(2), 289-301.

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: how attentional selection is shaped by ascribing intentions to others. *Plos one, 7*(9), e45391.

Wiese, E., Zwickel, J., & Müller, H. J. (2013). The importance of context information for the spatial specificity of gaze cueing. *Attention, Perception, & Psychophysics, 75*(5), 967-982.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592-598.

Wykowska, A., & Schubö, A. (2010). On the temporal relation of top-down and bottom-up mechanisms during guidance of attention. *Journal of Cognitive Neuroscience, 22*(4), 640-654.

Wykowska, A., & Schubö, A. (2011). Irrelevant singletons in visual search do not capture attention but can produce nonspatial filtering costs. *Journal of Cognitive Neuroscience, 23*(3), 645-660.

Wykowska, A., Wiese, E., Prosser, A., & Müller, H. J. (2014). Beliefs about the minds of others influence how we process sensory information. *Plos one, 9*(4).

Yerdon, V. A., Marlowe, T. A., Volante, W. G., Li, S., & Hancock, P. A. (2017). Investigating cross-cultural differences in trust levels of automotive automation. In S. Schatz S., M. Hoffman (Eds.), *Advances in Cross-Cultural Decision Making* (pp. 183-194). Cham, Switzerland: Springer International.

Złotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2016). Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy. *Paladyn, Journal of Behavioral Robotics, 7*(1).