On the Profoundness and Preconditions of Social Responses towards Social Robots

Experimental investigations using indirect measurement techniques

Inaugural-Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
im Fach Psychologie
der Fakultät für Psychologie und Sportwissenschaft
der Universität Bielefeld

vorgelegt von
Nina Riether
Bielefeld, Dezember 2013

Table of Contents

Acknowledgments

After hours spent in front of pictures and videos of robots, after numerous challenges of trying to get a real robot to look and act like it was supposed to, after countless days of measuring, analyzing, and interpreting people's behavior towards robots, and after the final stage of putting this research into writing, I fear I came very close to what might have looked, in fact, like a *social robot*. Despite this temporary shortcoming, I want to express gratitude to a number of people for their support and guidance throughout the various stages of this dissertation project. They all have had a significant influence on the completion of this work. Perhaps even more importantly, they helped me to maintain my sanity and humanity.

I want to express my gratitude to my supervisor Gernot Horstmann, whose academic and intellectual advice has always had a fruitful impact on the different stages of this work. His support has been invaluable. Further, I would like to deeply thank Britta Wrede and Franz Kummert for the academic freedom and the abundant support they provided at every step. I would particularly like to thank Johanna Kißler for her help and support without complications on short notice. Many thanks to all of you for your time and willingness to engage in the various challenges of this work with me.

Without the help of my colleagues at the applied informatics group, research assistants, and graduate students, the conduct and completion of this work would simply have been impossible. The sincerest thanks to Frank Hegel, Florian Berner, Maikel Linke, Michael Görlich, Andreas Kipp, Katja Schnatwinkel, Hendrik Stellmacher, Iris Sossalla, Lena Krönung, Jörn Weitz, Karola Hüppmeier, Iris Franzke, and Melissa Preuße. Without your effort, support, and advice these studies would have never found their way into the lab and, eventually, on paper. Further thanks go to my colleagues Sebastian Schneider, Ingmar Berger, Torben Töniges, and Luise Süssenbach for their social support, the discussions over coffee, and the necessary distractions from the daily routine.

I am very grateful to my former supervisors and mentors Katrin Lübke, Bettina Pause, Raoul Bell, and Axel Buchner, whom all helped to guide my first steps into research. During my dissertation years I profited greatly from your mentorship in research philosophy, theoretical and practical advice, indispensable programming and laboratory skills, as well as the many other lessons learned throughout our shared research experiences.

My profound gratitude goes to my friends. Every one of you contributed your fair share to the completion of this work. I cannot express how much I appreciate your practical help, your advice, your ability to cheer me up in difficult times, your questioning of my goals, and your will-

ingness to teach me new things. Thank you for being in my life, for keeping my eyes and ears open to the world, and for accompanying me on this journey!

Further, I want to express my great and deep gratitude to my parents and their unconditional love and relentless support. Thank you for challenging me, teaching me, and enabling me to gather my own experiences, while tirelessly providing moral, emotional, and motivational guidance. I fear I can never appropriately compensate all that you have done for me.

Ultimately, my profoundest gratitude goes to Cécile. I have the highest respect for you both personally and professionally. There are no words to describe my appreciation of your inexhaustible help, your precious advice, and your incredible patience. I am in awe of your forgiving generosity as well as your ability to perfectly dose the amounts of motivation, criticism, sympathy, and professional input needed to accompany me these last couple of years. I am forever indebted to you.

This work was supported by the German Aerospace Center (DLR, support code: 50RA1023) with funds from the Federal Ministry of Economics and Technology (BMBF) on the basis of a decision of the German Bundestag.

Abstract

Considering the future challenges of our ageing society and the already steadily increasing number of service robots for personal, domestic, and professional use, we will soon be faced with social robots as part of our daily lives. However, surprisingly, it is still unclear what societal role these non-human agents will play, and how exactly they should behave and interact with us. Will we be able to accept them as part of our social networks and be able to treat them like our human conspecifics? To answer these questions, we have to explore the basis of our socio-emotional behavior towards non-human agents, the underlying mechanisms, and boundary conditions. However, as yet, we still have not established a full understanding of how fundamentally social our reactions towards non-human agents are, how comparable they are to responses towards humans, and under what preconditions they occur. The present work was designed to examine the alleged profoundness as well as relevant preconditions of supposedly automatic social responses towards non-human agents using indirect measurement techniques. As previous literature has shown, it appears quite plausible that overt social responses found in human-human interaction may be transferable to human-robot interaction as well (e.g. Bartneck & Hu, 2008; Eyssel & Hegel, 2012; Kahn, Kanda, Ishiguro, Gill, et al., 2012; Klamer & Allouch, 2010; Kuchenbrandt, Eyssel, Bobinger, & Neufeld, 2011; von der Pütten, Krämer, & Eimler, 2011). Indeed, as far as adequate reference cues, e.g. humanlike appearance, are available, these mechanisms seem to work mindlessly and automatically. However, the concrete underlying processes behind these alleged social responses are not enlarged upon in great detail thus far. A closer look at subtle, low-level social and emotional reactions towards humans versus robots reveals that the mechanisms behind them are, in fact, less clear and potentially not truly equitable for human and robotic stimuli (Chaminade et al., 2010; Dubal, Foucher, Jouvent, & Nadel, 2011; Rosenthal-von der Pütten et al., 2013).

The application of experimental psychological methodology allowed for an investigation of the profoundness of these subtle, low-level responses by means of both indirect and unobtrusive measurements. This was thought to enable high standardization and conservativeness, while further ruling out confounding variables like demand characteristics, situational awareness, and self-report biases. Four studies were conducted that implemented these objectives. The first two studies utilized an applied setting, the social facilitation paradigm (see Aiello & Douthitt, 2001; Guérin, 1993; Zajonc, 1980). In a mere presence design, that is, the presence of a quiet, non-engaging spectator, the social influence of a present robot was assessed via performance-based measures on cognitive and motoric tasks. Not only was the mere presence of an active social robot able to elicit social facilitation effects, the performance effects were also comparable to those of human presence. Further, as the second study showed, these effects even sustained when the robot was obviously switched off.

These first two studies showed that the social influence of social robots indeed appears extremely profound and underlines the approach of mindless social responses elicited by basal social cues (Nass & Moon, 2000). Other explanatory concepts like demand characteristics, overt and conscious application of socially learned responses, superficial behavior adaptations, or reactions actually directed at the human behind the machine thus can most likely be ruled out, as the responses in these studies were measured indirectly and not via actual interaction.

To further examine the boundary conditions of these allegedly profound and low-level social reactions and to address the key factors of unfamiliarity, novelty, and humanlike appearance on a microlevel, two subsequent studies were executed that utilized the facial mimicry paradigm. Facial mimicry describes the rapid, unconscious motoric matching of a perceived facial expression, not necessarily visible by plain eye, but measurable via electromyographic activity over corresponding facial muscle sites (Cacioppo, Petty, Losch, & Kim, 1986; Dimberg & Thunberg, 1998). Additionally, these investigations allow further insight into the functionality of facial mimicry responses concerning the debate on the supposedly unspecific automaticity of these reactions versus them being a specific communicative tool particularly sensitive to human faces.

Results of the third study revealed that participants produced significantly less corresponding facial mimicry responses towards robotic faces compared to either human or schematic face types, and that this was especially apparent for happy facial expressions. Thus, additional aspects apart from mere humanlike appearance cues or a face-like configuration have to be met in order to evoke comparable socio-emotional responses towards humans and robots on the basal microlevel. Stimulus ambivalence, possibly caused by stimulus unfamiliarity, as well as specific physical stimulus characteristics of the happy faces, concretely, the intensity of the displayed smile, are discussed as potential influence factors for these findings. The final study thus examined whether the previous differences between facial mimicry responses towards human and robotic faces could be mitigated by familiarization as well as altered expression intensity of the happy faces. Indeed, real-world familiarization through actual interaction with a social robot in the past together with more pronounced smiles of the robotic faces evoked comparable mimicry responses towards robotic as towards human faces. Hence, the formation of a general mental model of robots through previous actual robotic interaction along with pronounced expression intensity of the displayed smiles appeared as the needed link in order to produce appropriate and comparable basal socio-emotional responses towards robots.

Stimulus ambivalence is discussed as potent moderator of the obtained results, both in the social facilitation as well as the facial mimicry paradigm. Devoid of a mental model of robots in general, a presented robot will apparently represent an ambivalent stimulus, triggering subsequent mechanisms like monitoring responses, heightened arousal, distraction, increased situational awareness, etc. As the social facilitation studies revealed, thereby, even a switched-off robot was able to impair human performance. Additionally, this may even affect socio-emotional responses on the most basal level – fast, invisible, motoric facial mimicking responses.

Taken together, on a behavioral macrolevel, as the social facilitation studies showed, a face-like configuration or other humanlike appearance cues apparently sufficed as fundamental initiation for social responses towards non-human entities, however, looking at low-level reactions, which were measured indirectly, humanlikeness alone was not sufficient for the emergence of congruent facial mimicry responses. Rather, the interaction of specific preconditions, that is, real-world familiarization and more pronounced stimulus intensity, presented the crucial pivot regarding the underlying mechanisms behind the social responses. Thus, while going along with the prominent approach of Nass and colleagues (e.g. Nass & Moon, 2000; Reeves & Nass, 1996) concerning the apparent "mindlessness" of social responses towards non-human entities, this work depicts a more nuanced approach, as obviously a minimal cue alone was not sufficient in activating low-level social response schemes towards robots resulting in fully comparable reactions as towards human stimuli (contrasting Epley, Akalis, Waytz, & Cacioppo, 2008; Nass, Steuer, & Tauber, 1994; Nowak & Biocca, 2003). Whereas humanlikeness may be indeed essential regarding the initial activation of social response schemes (Epley, Waytz, Akalis, & Cacioppo, 2008), other factors, like stimulus familiarity or stimulus intensity, might overlay these and alter the manifestation of concrete social and emotional responses.

As the findings of the present work appear fruitful both for the fields of social robotics as well as social and biological psychology, additionally, the current results are discussed in light of the implications for the actual application of social robots in various domains and contemplates on the usage of robots and other non-human agents as research vehicle. Further, the present findings are integrated with the explanatory concepts of fluency (Reber, Winkielman, & Schwarz, 1998; Winkielman & Cacioppo, 2001; Winkielman, Schwarz, & Reber, 2000) and, specifically regarding the facial mimicry studies, the neuropsychological basis and functionality of the respective muscle responses (Larsen, Norris, & Cacioppo, 2003; Rinn, 1984). Additionally, they may allow for a potential distinction of the effects of humanlikeness from anthropomorphization responses (Epley, Waytz, & Cacioppo, 2007).

To conclude, the application of indirect and unobtrusive forms of measurement proved a sensible and successful methodology in exploring the boundary conditions of the apparent automaticity and profoundness of social responses towards non-human agents.

1. General Introduction

"At bottom, robotics is about us. It is the discipline of emulating our lives, of wondering how we work" (Rod Grupen in Hapgood, 2008).

Nowadays, while moving through our day-to-day live, we are confronted with myriads of little electronic helpers and diverse forms of artificial agents. Interaction with electronic equipment happens to be quite natural today, ever since computers became widely available for home and office use. Online services, smartphone applications, immersive entertainment equipment, and many other electronic devices already make heavy use of artificial humanlike agents. As Krämer (2008) noted, it seems not to be question whether these social agents will find their way into our daily lives, but rather a matter of how exactly we will want this to look like.

While we have gotten more and more familiar with virtual humanlike agents, nowadays, so-called social robots emerge as particularly interesting. They are defined as physically embodied agents, usually equipped with a humanlike appearance and basal abilities of interaction, communication, and cooperation, thus taking human-agent interaction from a level of on-screen- (e.g. virtual vocabulary coach) or assistance-tool- (e.g. robotic vacuum cleaner) applications into actual physical social environments with both additional challenges and chances. The scope of applications of social robots is incredibly diverse, ranging from fields of health care, work and domestic assistance, service activities to therapy, tutoring scenarios, and entertainment purposes (Dautenhahn, 2007; Kanda, Ishiguro, et al., 2004; Mataric, 2008; Tapus, Mataric, & Scassellati, 2007).

In 2012, the number of service robots sold for professional, personal, and domestic use already outnumbered the sales for industrial robots by the twentyfold and is expected to increase even stronger during the next couple of years (IFR Statistical Department, 2013). Specifically regarding the social and budgetary impact of our aging society, social robots have the potential to be especially beneficial. Imagine that in the near future, we might be treated by robotic nurses at hospitals or be taken care of in our own homes, have individual shopping assistants, and discuss our personal problems with a robot therapist. At work, a robotic aid helps us finish the latest project, whereas at home, children are playing a tutoring game with the robot nanny, while we plan the dinner menu and cleaning chores for the next couple of days with the household helper robot. We might even be able to give our distant grandparents a hug over the remote communications robot. As a social interface is regarded as universal interface (Breazeal, 2003; Reeves & Nass, 1996), these robots will probably share a humanlike morphology and further components to support the natural communication with humans, like facial expressions, body posture, gesture, and voice. Thereby, we will be able to interact intuitively, naturally, and enjoyably with our robot companions (Breazeal, 2003).

However, even though they will likely form an important part of our future society, their social role and functionality is still largely unclear. Will we treat social robots in the very same way we treat our conspecifics? Are we able to accept social robots as part of our social fabrics on a permanent basis? Do we have to apply the same moral standards to social robots as towards humans? Under what circumstances will we stop treating social robots like humans? Can we harm them, switch them off, tear them apart, and sell them without feeling bad? Will we be able to make them accountable for our feelings and emotions, our social reactions, our work performance? Put differently, how profound and "hard-wired", but also how far-reaching and flexible is our ability to treat non-human agents as humans? To answer these questions, we need a thorough understanding of the mechanisms and boundary conditions of social reactions towards social robots.

Indeed, for some time now, research on artificial humanlike entities has been going a step further than the classical evaluation of efficiency, effectiveness, and acceptance of interaction with these agents and started to include investigations of social effects of virtual agents and robots. The overall findings present the recurring evidence of, indeed, social responses towards these entities, apparently comparable with social reactions towards other humans (see next section; e.g. Epley et al., 2007; Eyssel & Kuchenbrandt, 2012; Hoffmann, Krämer, Lam-Chi, & Kopp, 2009; Krämer, 2008; Kuchenbrandt et al., 2011; Nass & Moon, 2000; von der Pütten, Krämer, Gratch, & Kang, 2010; Sproull, Subramani, Kiesler, Walker, & Waters, 1996).

However, one of the most intriguing questions in this line of research is the controversy whether these social responses are either superficial or even conscious adaptations based on for example the displayed humanlikeness or rather emerge as purely automatic and involuntary "hard-wired" reactions by virtue of the fundamental social nature of humans. This debate alludes to sensible reflections on how profound social reactions towards non-human entities truly are, and which factors play a vital part in initiating or influencing these. The exploration of the preconditions for social reactions towards non-human agents thus provides valuable insight into an essential question, bearing relevance both for basic and applied research – what are the minimum requirements for social reactions? Following this, once we identified the minimum requirements for social reactions, what are the implications for the conceptualization, design, and application of social robots?

Hence, apart from evaluating the effects of the application of social robots and giving empirically founded suggestions regarding the design and interactional content of human-agent interaction, experimental psychology can equally gain insights into the boundary conditions of social cognition by closely investigating the subtleties of social responses towards non-human entities. Integrating approaches both from biopsychology as well as social psychology appears as a reasonable and likewise rewarding, yet sparsely represented research approach in tackling the aforementioned questions. The following sections will comprise a brief overview on the research targeting the various social responses towards robots and other comparable non-human agents as well as an outline of potential underlying explanatory concepts.

Social responses towards non-human entities

First off, the current social agents and robots are not able to even closely resemble a human interaction partner regarding emotional expression, social behaviors, and communicative features. However, it is all the more interesting that these entities still evoke social reactions in their human counterparts. This social influence manifests itself in variously observed social behaviors towards computers, virtual agents, and robots.

In their extensive and prominent work Nass, Moon, Morkes, Kim, and Fogg (1997) as well as Nass and Moon (2000) showed that people even treated conventional computers as social entities and conclude that interactions with computers may be, in fact, fundamentally social. This was termed *media equation*. Stereotypization, utilization of social norms like politeness, basal rules of perception and evaluation, but also concepts like cooperation and group processes all seem to be applicable to human-computer- and, as subsequent research showed, human-agent-interaction (Bartneck & Hu, 2008; Hoffmann et al., 2009; Nass, Isbister, & Lee, 2000; von der Pütten et al., 2010).

For instance, human's automatic tendency to categorize others based on characteristics like gender or ethnicity and to accordingly define in-groups vs. out-groups has been shown to also pertain to robots (Eyssel & Hegel, 2012; Eyssel & Kuchenbrandt, 2012). This even holds true for arbitrarily formed groups, for example different color-marked teams (see *minimal group paradigm,* Tajfel, Billig, Bundy, & Flament, 1971), which then leads to comparable results regarding in-group cooperation and –evaluation (e.g. Kuchenbrandt et al., 2011; Nass, Fogg, & Moon, 1996). Remarkably, these social reactions seem to go in hand with additional attributions of other humanlike qualities and even personality traits towards the artificial entities (see Epley et al., 2007). Participants usually even report higher levels of credibility, authenticity, and trust towards humanlike agents (Sproull et al., 1996).

Some anecdotal evidence suggests that people show emotional responses and even attachment towards robots (Klamer & Allouch, 2010; von der Pütten et al., 2011). Indeed, participants argued with a robot that deceived them (Kahn, Kanda, Ishiguro, Gill, et al., 2012), and reacted physiologically and behaviorally calmer towards an affectively expressive versus neutral robot in a search-and-rescue-context (Bethel, 2009). Replications of Milgram's obedience experiments (Milgram, 1974) with robots yield anecdotal evidence that people seemingly felt hesitation and pity with the robot they had to electrically shock or destroy with a hammer, some even calling the experiment "inhumane" (Bartneck & Hu, 2008; Bartneck, Chioke, Menges, & Deckers, 2005). Accordingly, even complex concepts like empathy may be transferred to human-robot interaction, as research participants also felt pity for robots that repeatedly fell down a hill (Dautenhahn, 1998, as cited in Krämer, 2008) or were stuffed back into a closet before they could finish an ongoing game (Kahn, Kanda, Ishiguro, Freier, et al., 2012).

Evidence for emotional reactions towards social robots is not only found on subjective or behavioral, but also on a physiological level: For example, witnessing the torture of a robot does not only elicit subjectively reported empathic concern but is immediately quantifiable by an increase in physiological arousal (assessed via electrodermal activity and heart rate; Rosenthal-von der Pütten, Krämer, Hoffmann, Sobieraj, & Eimler, 2012). Additionally, Rosenthal-von der Pütten et al. (2013) found comparable neural activation patterns in limbic structures for stimuli depicting violent actions against a human as well as a robot.

Because neural and physiological responses are relatively immune to demand characteristics, self-report biases, and other potential confounds through direct interaction, they are especially valuable for the exploration of social and emotional reactions towards non-human entities. Neural and physiological responses allude to the very core of our perception and behavior. Given that the neural level is the biological essence for social interaction, identical neural activity towards robots and humans would indicate that these reactions towards non-human agents are indeed ultimately social and profound.

Yet, as will be presented in the following, the neural processes behind these responses appear not as equitable as it might seem. As social responses build on a neural architecture, this might subsequently affect the overall acceptance of non-human agents and the formation of stable, long-term interactions. We will most likely behave and feel differently around robots in our homes and workspaces that we can form positive, natural, or humanlike relationships with compared to others that we find ambiguous, tool-like, or that trigger deeply rooted avoidance behavior.

In particular, Rosenthal-von der Pütten et al. (2013, see above) discovered differences in neural activity that imply higher empathic concern for humans compared to robots. In a similar vein, Dubal et al. (2011) investigated whether our brains respond similarly to robotic emotional expressive stimuli as to human emotional expressions via event-related potentials. Their data showed that the encoding of emotional expressions apparently did not differ between human and robotic stimuli; only the processing of face-like properties was delayed and overall reduced for robotic emotional expressions. Thus, even though human and robotic faces may differ in their representation of humanlikeness, early brain processing of emotional expressions seems indeed comparable.

In an fMRI study, Chaminade et al. (2010) explored potential differences in the neural processes involved in reading the emotions displayed by either a human or a robot. They found increased responses in areas relevant for visual processing for robotic compared to human stimuli. Additionally, reduced activity in areas being potentially involved in supposed mirror processes (Buccino et al., 2001; Chaminade & Decety, 2001; Iacoboni et al., 1999; Rizzolatti & Craighero, 2004) for robotic emotional stimuli suggests lower resonance with the robot as opposed to the human. However, explicit instructions to concentrate on the emotional content of the stimuli led to significant increases in neural resonance towards the robotic stimuli. The authors discuss familiarity and humanlike appearance as potential mediating factors for their results.

Thus, whereas some of the fundamental processes behind these profound responses appear similar for human and robotic stimuli (e.g. the P1 component in ERP responses, Dubal et al., 2011; comparable limbic activation patterns after violence against a human and a robot, Rosenthal-von der Pütten et al., 2013), they do not present as completely interchangeable (more empathy towards humans, Rosenthal-von der Pütten et al., 2013; lower resonance for robotic stimuli when no explicit instruction is given, Chaminade et al., 2010), and potential influence factors are brought up for discussion once already.

Explanatory concepts

While the mere social and emotional reactions to non-human entities are already remarkable themselves, research rather recently began pursuing to unravel their possible explanatory pathways as well as implications for underlying psychological processes. Revisiting the concept of the media equation, Nass and colleagues (e.g. see Reeves & Nass, 1996) propose their concept of *ethopoeia* as underlying factor for social and emotional reactions to non-human entities, which they term the automatic and unconscious social reactions towards artificial humanlike entities, despite the knowledge that these entities are, in fact, not human and do not warrant human treatment. On grounds of the fundamental social nature of human behavior, the slightest social reference cue would unconsciously trigger deeply rooted social reaction schemes (Nass & Moon, 2000; also see Langer, 1992, regarding mindless social behavior in response to social contextual cues) under the premise of the ubiquity of social norms on the one hand and the perceived humanlikeness (in terms of a representativeness heuristic) of the entity on the other hand (Sundar & Nass, 2000). Following this approach, simply equipping a robot with a pair of eyes for example thus should suffice in triggering relevant social responses towards the agent in a specific social situation.

A different theory adds demand characteristics of the respective situation for consideration as crucial determinants for social reactions towards non-human agents, which are thought to be superficial adaptations rather than deeply rooted behavioral tendencies (Kiesler & Sproull, 1997). A social reaction towards for instance a humanlike robot would always be an "as if"-behavior and not a real attribution of humanlikeness. According to this view, for example we will always greet and thank the robot server at a restaurant, simply because we superimpose our social scripts for specific social situations on the current interaction with the machine, fully aware that our counterpart is a machine. Because we retrieve the specific script for behaving at a restaurant and the machine acts like a server, we will respond in fashion of a superficial adaptation, that is, we will treat the machine as if it were a server. However, this specific behavior pattern does not require that we perceive the machine as a fundamentally social and coequal interaction partner. This implies that the acceptance of social robots might naturally never cross the line towards authentic humanlike treatment and respective reactions.

Juxtaposing these explanatory approaches, the subtlety and largely unconscious nature of the measured social reactions (usually, participants are not even aware of themselves behaving socially towards the non-human agent) speak in favor of the first line of thought, highlighting deeply rooted social reaction schemes (Nass & Moon, 2000). Still, in order to clarify the controversy between superficial adaptations on the one hand and deeply embedded automatic reactions on the other, there is a need for studies exploring subtler, delicate, low-level social responses.

A recurring approach, potentially reconciling the previous claims within a psychological basis, stems from the prominent concept of anthropomorphism regarding social reactions towards robots or artificial agents. Anthropomorphism describes the common tendency of imposing humanlike characteristics, character traits, intentions, motivations, and emotions on non-human entities (Epley et al., 2007). Epley et al. propose a three-factor model of anthropomorphism that includes the following three key psychological determinants: accessibility and applicability of agent knowledge, motivation to understand and explain the behavior of other agents, and the need or desire for affiliation and social contact. The first factor alludes to the general availability and representativeness of knowledge about the self or other humans that serves as a basis for inductive reasoning about non-human entities. That is, as long as our knowledge about the category "human" is more readily accessible and more detailed than that about the category "non-human", we will apply our knowledge of the former to the latter. When a robot looks human and we have little knowledge about robots, we supposedly are likely to apply our knowledge of humans to the robot. The second factor, effectance motivation, aims at the dispositional need for closure and desire for control when interacting with ones environment. Anthropomorphization thus offers a vehicle to reduce uncertainty and deal with unpredictability when reasoning about or interacting with non-human agents. Hence, in cases of uncertainty, anthropomorphization may serve as practical resource for understanding non-human agents by attributing humanlike characteristics, intentions, and motivations to them, regardless of their accuracy. When the robot's actions are unpredictable, we supposedly likely infer from what we would expect another human to behave in this situation. The last factor, sociality motivation, describes the need for social connection. Hence, when people feel lonely or socially disconnected, e.g. consider patients at a nursing home, they are apparently more likely to anthropomorphize. Indeed, research has shown that these three factors seem to influence anthropomorphic inferences and subsequent social responses towards non-human entities (Epley et al., 2007; Epley, Akalis, et al., 2008; Epley, Waytz, et al., 2008; Eyssel & Kuchenbrandt, 2012; Eyssel, Kuchenbrandt, & Bobinger, 2011; Waytz, Cacioppo, & Epley, 2010).[1]

As several humanlike characteristics evidently influence the degree of anthropomorphic inferences (Epley, Akalis, et al., 2008; Eyssel & Hegel, 2012; Eyssel, Hegel, Horstmann, & Wagner, 2010; Nowak & Biocca, 2003), accordingly, the essence of anthropomorphization is perceived humanlikeness (Epley, Waytz, et al., 2008). Indeed, perceived humanlikeness seems to be a very potent cue for eliciting social responses and thus might serve as prime candidate for moderating social reactions towards social robots. Clearly, social robots, unlike virtual agents, possess an inherently human feature – a physical "body". Accordingly, social robots outperform virtual agents on various measures, like attributed social influence and experienced positivity of attitudes towards these entities (Powers, Kiesler, Fussell, & Torrey, 2007). It is further demonstrated that facial features and head dimension significantly determine the amount of perceived humanlikeness (DiSalvo, Gemperle, Forlizzi, & Kiesler, 2002). This is in line with the overall biological significance of faces in visual perception. Given that faces are extraordinarily potent social and emotional stimuli, humans are supposedly equipped with a functional sensitivity for faces, face-like cues, and facial expressions (Haxby, Hoffman, & Gobbini, 2000; Öhman, 2002).

More specifically, *emotional* faces seem particularly ecologically relevant in that they elicit specific cortical responses in the observer (e.g. see Haxby et al., 2000, for a short overview). Accordingly, EEG and fMRI findings of stronger perceptual activations in visual regions[2] for emotional

---

[1] Along with the theory of anthropomorphization, concepts like the *intentional stance* thesis (Dennett, 1989; 1996) or *theory of mind* approaches (Call & Tomasello, 2008; Premack & Premack, 1995; Premack & Woodruff, 1978) also claim that people tend to attribute human intentionality to all possible sorts of entities. Whereas this could be caused by overgeneralization processes (Dennett, 1996), another option might be that differentiation categories between "human" and "non-human" might not be salient or accessible enough in human-agent interaction (Nowak & Biocca, 2003), thus leading to social responses towards non-human entities. This again is in line with the first influencing factor *elicited agent knowledge* of Epley et al.'s (2007) model of anthropomorphism.

[2] Several distributed neural structures are involved in the processing of visual social cues and faces, like the inferior occipital gyri, superior temporal sulcus, lateral fusiform gyrus, the amygdala (being especially relevant as soon as emotional content is involved), insula and limbic system as well as the orbiftorontal cortex (Adolphs, 2004; Allison, Puce, & McCarthy, 2000; Haxby et al., 2000; Öhmann, 2002).

compared to neutral faces support this notion (Adolphs, 2004; Haxby et al., 2000; Morris et al., 1998; Nakamura et al., 1999; Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998; Vuilleumier, 2005; Vuilleumier, Armony, Driver, & Dolan, 2001). This seems to not only hold true for human but schematic faces as well (Britton, Shin, Barrett, Rauch, & Wright, 2008; Eger, Jedynak, Iwaki, & Skrandies, 2003). Thus, even clearly non-human emotional stimuli only including basic face-like cues seem able to elicit comparable brain responses in the specific cortical networks. This again links nicely to the EEG results from Dubal et al. (2011) already mentioned above. Even though the brain reacted differently regarding the face-like representation of human versus robotic stimuli, the responses related to the emotional content of the facial expression were indeed comparable.

Drawing an interim conclusion, humanlikeness appears as key determinant of social responses towards non-human entities. Does this imply that the more humanlike the robot is designed, the more social responses it will elicit? Will a more humanlike robot be naturally met with higher acceptance? Apparently, the relation is not as simple, as the following sections will describe.

An ongoing issue, thus, is whether the relationship between humanlikeness of social robots and evoked social responses is linear. According to Nass and colleagues (Nass et al., 1994; 1997; Nass & Moon, 2000) even the most primitive cues can turn a computer into a “social actor”. Nass and colleagues argue that social responses towards non-human agents not necessarily have to be cued by a “rich human presentation” of the artificial entity (Nass et al., 1994), in terms of a face for instance, but rather that subtle, even primitive cues suffice in generating social responses. Kiesler and Sproull (1997) on the other hand regard the quantity of humanlike attributes as critical for evoking appropriate social responses.

Closely related to this debate is the *uncanny valley* concept (Mori, 1970). Originated as a thought experiment, it has been shown that increased humanlikeness only linearly affects familiarity, acceptance, and social responses up to a certain degree: Familiarity and acceptance slowly grow with rising humanlikeness, but rapidly decrease again (and even fall in the “eerie” spectrum) shortly before complete humanlikeness is reached. Prominent explanations pose norm violations (MacDorman & Ishiguro, 2006a; Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012) and conflicting social cues (Moore, 2012) as causal for these effects. Regarding the first, an artificial entity looking almost human will be met with the same mental models and expectation standards as a real human being would. The notable shortcomings and imperfections of the agent will thus not be held against the artificial-human-standard but rather the human-standard, resulting in a feeling of strangeness or even creepiness (MacDorman & Ishiguro, 2006a; 2006b). Second, contradictory social cues such as ambiguous signals of category membership (e.g. faces neither clearly belonging to a human nor to a non-human category) may produce subjective eeriness (Burleigh, Schoenherr, & Lacroix, 2013).

Taken together, humanlikeness (within a certain range, as the uncanny valley findings suggest) seems to foster social responses towards artificial agents, if not initiates them. However, it is still unclear on what level humanlikeness affects these social responses. Does humanlikeness operate on a low-level purely automatic and "mindless" account? Or does perceived humanlikeness rather lead to superficial conscious anthropomorphization responses?

Given that perceived humanlikeness has the potential to serve as the basic key mechanism for social responses towards non-human entities, present research has mostly neglected moderating variables such as novelty and familiarization effects. The time-sensitivity of social reactions towards non-human entities is still largely unclear, mainly due to difficulties in the conduction of both technically and experimentally reliable long-term studies and the related financial constraints. Novelty effects pose social reactions towards non-human agents to only occur as long as the agent is new and unfamiliar. Accordingly, as an appropriate mental model for the non-human counterpart is missing at first, well-known social scripts from human-human interaction are utilized, which should be gradually replaced by an adapted mental model and behavior script for the non-human agent. However, if any kinds of social cues persistently lead to social reactions as Nass and colleagues posit (see e.g. Nass & Moon, 2000), social behavior towards non-human entities should occur on a long-term basis.

Present research questions

Drawing a conclusion from previous literature, it appears plausible that overt social responses found in human-human interaction seem indeed transferable to human-robot interaction, as far as adequate reference cues are available. Additionally, these routines appear to work mindlessly and automatically. Succeeding research in the fields of social robotics and human-robot interaction is usually spreading apart the scope of these social reactions across the many diverse application fields. However, the exact underlying processes behind these responses are not enlarged upon in great detail thus far. Yet, the closer we look at more subtle, low-level social and emotional responses, accurate comparisons between human and robotic stimuli firstly are represented rather scarcely in the literature and secondly yield mixed results concerning the equalization of the underlying processes. Despite indicating *comparable* pathways, the low-level responses do not appear to be completely and interchangeably *equitable*. Are social responses towards non-human entities not as profound as prior research has suggested? To investigate the alleged profoundness of social reactions towards non-human agents, we need to focus on the basal mechanisms and boundary conditions of these social and emotional responses.

Merging the application of experimental psychological methodology in the investigation of social and emotional effects of social robots with ancillary exploration of the minimum requirements for social responses in general, two study objectives appear appealing. First, it seems sensible to fix the measurement of social responses to indirect and unobtrusive forms in order to allow for high standardization and to discard potential factors like demand characteristics or heightened situational awareness. Second, paradigms involving no direct interaction are convenient as this inherently enables the reduction of social responses to their more subtle, low-level components, thus being maximal conservative concerning the profoundness of these reactions and ruling out confounding with specific interactional characteristics.

Two paradigms that fulfill these objectives are the social facilitation and facial mimicry paradigm. Social facilitation describes the influence of social presence on human performance and is one of the oldest paradigms of social psychology (for an extensive review see Aiello & Douthitt, 2001; Guérin, 1993). Specifically, the influence of social presence on task performance has been established to be dependent on task difficulty – for easy or well-learned tasks, the mere presence of another person facilitates performance (hence the terminology), whereas performance decreases on difficult or new tasks, compared to an alone control condition. This paradigm is particularly neat for investigating social responses towards non-human entities, as it does not require a direct social interaction with the entity, but enables to employ an indirect yet very precise measure, namely human cognitive and motoric performance, of the influence of social presence. Addi-tionally, the experimental setting seems highly relevant regarding the future application of artificial agents and, specifically for this work, robots in workplace scenarios. Extending the investigation of the basic mere presence social facilitation paradigm to allude to possible influencing factors, it appears sensible to examine whether potential social facilitation effects still occur when the robot is obviously switched off. Thus, the following research questions were investigated in two subsequent studies:

1) Is the mere presence of a social robot able to elicit social facilitation effects?
2) Are these potential effects comparable to those of human presence?
3) Will potential effects be dependent on perceived animacy, that is, whether the robot is active versus switched off?

Whereas the social facilitation studies aim at capturing the effects of social influence in a very basic and conservative way through an applied setting, the second paradigm, facial mimicry, takes the investigation of social responses to the microlevel, which makes it particularly suitable for examining the minimum requirements of these responses. Facial mimicry terms the rapid and seemingly unconscious mimicking of the facial expressions of our counterpart, occurring as fast as 300-400 ms after stimulus onset (Dimberg & Thunberg, 1998). As this is not necessarily visible by plain eye, these responses can be measured via electromyography of corresponding facial muscle activity (Cacioppo et al., 1986). This paradigm offers the unobtrusive investigation of subtle, delicate emotional responses towards non-human entities, thus adding a novel measurement technique and concomitant theoretical foundation of these responses as low-level, socio-biological reactions to the few neuropsychological and physiological studies in this field.

Comparing the amount of mimicking of facial expressions of robotic versus more human-like and also more schematic faces can both illustrate the influence of specific appearance factors for the formation of instant social-emotional responses towards non-human entities as well as potentially disentangle the debate whether facial mimicry itself is an unspecific, automatic, non-conscious response or rather a purposeful communication tool that may be particularly sensitive to human faces. Additionally, facial mimicry responses seem especially suitable to explore the potential effects of manipulation of facial cues, for instance in terms of stimulus intensity or facial configuration, constituting promising results both regarding potential influence factors for social responses towards social robots as well as the neuropsychological foundation of facial electromyographic responses in general. Furthermore, previously specified as unaccounted moderating factor, the facial mimicry paradigm offers a smaller scale approach to address novelty and familiarity effects apart from conducting long-term studies, by considering response facilitation due to different forms of enhanced stimulus familiarization. Two studies were implemented in order to explore the following questions:

1) Will facial mimicry differentiate between human, robotic, and schematic faces?
2) Will potential differences between face types be diminished by altered stimulus characteristics like stimulus intensity?
3) Will familiarization diminish potential differences between face types?

2. Social Facilitation with Social Robots[3]

---

[3] Parts of this chapter have been published as Riether, N., Hegel, F., Wrede, B., & Horstmann, G. (2012). Social Facilitation with Social Robots? *Proceedings of the 7th International Conference on Human-Robot Interaction*, 41-48.

2.1 Theoretical and Empirical Background

The social facilitation paradigm offers an elegant way to examine basal components of social responses towards non-human entities indirectly and unobtrusively, thus being very conservative regarding the profoundness of these responses, while at the same time utilizing a relevant applied setting – the influence of social presence on human task performance. Considering the various existing application domains (e.g. Dautenhahn, 2007; Kanda, Hirano, et al., 2004; Mataric, 2008; Tapus et al., 2007) and the future usage of social robots in workplace scenarios, where robotic support may appear highly useful, but where human performance is still of indispensable importance, it is all the more essential to examine a pertinent question: What are the effects of mere robotic presence on human performance?

The social facilitation effect is one of the oldest paradigms in social psychology (Triplett, 1898) and describes differential performance variations depending on task difficulty in the presence of another person. Specifically, performance on easy or well-learned tasks are facilitated when another person is present, whereas performance on complex or new tasks is impaired in the presence of another individual, compared to an alone condition. After prominent experimental investigations by Zajonc (1965; 1980) and others (Baumeister, 1982; Berger et al., 1981; Bond, 1982; Carver & Scheier, 1978; Cottrell, 1972; Duval & Wicklund, 1972; Easterbrook, 1959; Jones & Gerard, 1967; Kushnir & Duncan, 1978; Sanders & Baron, 1975; Zajonc, Heingartner, & Herman, 1969), various explanatory theories evolved, which can be roughly grouped into the three main categories (see Guérin, 1993) of drive- or arousal-theories, social comparison theories, and thirdly, cognitive process theories.

Arousal-theories emphasize Hull-Spence drive theory (Spence, 1956) as explanatory mechanism. Zajonc (1965; 1980), relating to drive theory, proposed that the presence of others would increase individual drive or arousal levels, which in return enhanced the probability, or reaction potential, for dominant responses. These usually are correct or fast responses on simple or well-learned tasks and, accordingly, incorrect or slow responses on complex or new tasks. This was thought to explain the observed performance enhancements and decreases.

Cognitive process theories pose attentional factors, like physical distraction (Kushnir & Duncan, 1978), attention conflicts, shared attention (Jones & Gerard, 1967; Sanders & Baron, 1975), or restricted focus of attention (Easterbrook, 1959) as prime explanations for the obtained performance effects. Basically, these mechanisms naturally play a part in all explanatory theories.

Social comparison processes build another explanatory approach, including classical concepts like evaluation apprehension, self-awareness, and impression management (Baumeister, 1982; Bond, 1982). However, other than cognitive process or arousal theories, they embrace a deliberate choice of socially acceptable behavior and likely an inhibition of socially undesirable behavior in order to meet specific situational requirements. Thus, subjective matching of actual and ideal behavior, concern about the own performance, worry about being evaluated, comparison with social norms and standards (e.g. Carver & Scheier, 1978; Cottrell, 1972; Duval & Wicklund, 1972), but also behaviors like explicit suppression of rehearsal strategies in front of an observer (Berger et al., 1981) may all lead to the predicted effects. Obviously, these factors also rely on mechanisms like restricted focus of attention or other attentional conflicts as well as heightened arousal.

Taken together, none of the three explanatory concepts seems solely applicable in accounting for all the specific results in the field of social facilitation research; yet, all theories appear to roughly predict the same findings of the basic paradigm. Consequently, some authors suggest an interaction of the different proposed mechanisms (Aiello & Douthitt, 2001; Guérin, 1993) in a larger network of social responses, next to for instance social loafing, cooperation, and competition. Social facilitation appears as a complex network of cognitive and biological functions, social standards, and individual learning history. It seems substantially dependent on the specific situational characteristics which explanatory mechanism applies to what degree.

Dashiell (1930) identified various types of presence and specified the most basic being the mere presence of a quiet, non-engaging spectator. According to Zajonc (1980), mere presence should suffice in obtaining social facilitation effects, as the mere presence of another person could already lead to enhanced alertness, thus inducing arousal. Although the presence of an observer could also induce heightened evaluation potential, after Zajonc, evaluation apprehension would not be strictly necessary. Additionally, Zajonc claimed social stimuli to be highly unpredictable (Zajonc, 1980), which in turn could cause heightened alertness (Norman, 1980) and "hard-wired" monitoring responses (Lynn, 1966; see Guérin, 1993, for extensive review), again ensuing enhanced arousal.

Overall, all theories indicate the necessity of perceived co-presence of another individual in order to obtain the expected social facilitation effects. As this paradigm does not involve direct interaction for the unobtrusive measurement of social responses, it is of particular interest whether a social robot will also be able to elicit social facilitation responses in humans utilizing very precise measures of cognitive and motoric performance. Regarding potential comparable effects to those of human presence, this could draw several implications for the application of robots in workplace-scenarios, for instance as tutor, task companion, or assistance tool, as their presence might generate unwanted effects, depending on the task specifics. Robotic presence might impair human task performance when the task is new or complex, but enhance performance for simple or well-learned tasks. Additionally, if robotic presence might lead to performance impairments, will simply switching it off terminate its social influence? That way, potential confounds of novelty, or unfamiliarity with animacy can be neatly addressed. However, none of these questions has been empirically investigated as of today.

Regarding related works, a couple of studies have already explored whether social facilitation effects are transferable to virtual agents (Rickenberg & Reeves, 2000; Walker, Sproull, & Subramani, 1994; Zanbaka, Ulinski, Goolkasian, & Hodges, 2004). Unfortunately, difficulties regarding the methodology and dependent measures as well as problems due to ceiling effects on simple tasks have lead to mixed results and complicate the comparability to the classical paradigm. Additionally, manipulation of varying task difficulty was not experimentally controlled for (see Krämer, 2008).

Yet, several studies that applied thorough experimental methodology evidently succeeded in replicating the social facilitation effects (Hoyt, Blascovich, & Swinth, 2003; Park & Catrambone, 2007; Zanbaka, Ulinski, Goolkasian, & Hodges, 2007) for the presence of virtual agents. Regarding the presence of social robots (Bartneck, 2003; Woods, Dautenhahn, & Kaouri, 2005), however, the obtained findings are vastly confounded by methodological flaws (no alone control group, no experimental manipulation of task difficulty were used).

As no experimental study exploring mere presence effects of social robots on human task performance has been conducted yet, the following two investigations aimed at filling this gap. Encouraging findings from the field of virtual agents in this domain as well as the vast body of research regarding overt social reactions towards non-human entities suggest comparable results to that of human presence. However, physical embodiment might even lead to stronger results than those obtained with virtual agents.

Thus, in the first study, it was expected that the mere presence of a social robot as well as the presence of a human should be able to elicit the predicted social facilitation effects compared to an alone condition. Another undirected hypothesis targeted the potential comparability of human to robotic presence, which still is unaccounted for as of today. Hence, the performance on four cognitive and motoric tasks, which were administered both in an easy and complex version, was determined for three presence groups (alone vs. robotic presence vs. human presence). By using subtle, indirect, yet precise measures of human performance, this first study offers a suitable instrument to explore the depth and transferability of basal social responses towards non-human entities.

Consequently, the second study was conducted to further enlarge upon the profoundness of these responses by examining whether potential robotic presence effects still prevailed when the robot was obviously switched off. As this is of apparent importance regarding the practical application of these findings, it additionally embraces contemplation on the potential influencing factors of novelty, animacy, and unpredictability for the explanatory concepts of social responses towards non-human entities in general. For the second study, the task setup of Study 1 was utilized, however, the three presence conditions now consisted of an alone control group, a second group with an active robot (comparable to the robotic presence condition of Study 1) as well as a third group with the robot evidently switched off.

2.2 Study 1 – Replication of the Social Facilitation Paradigm with Social Robots

2.2.1 Methods

2.2.1.1 Participants and design

One hundred six right-handed participants (78 women, no gender differences between groups) with a mean age of 23.31 years (*SD* = 2.99 years) were randomly assigned to one of three groups resulting from a 3 (*presence type:* alone, human confederate present, robot present) × 2 (*task difficulty:* easy, complex) mixed-factorial design. Presence type served as between-subjects factor whereas task difficulty was manipulated within-subjects. The performance on four different tasks of cognitive and motoric nature (*anagram solving*, *numerical distance*, *finger tapping,* and *motoric tracking*, see 2.2.1.2) served as dependent variable. The tasks were administered computerized both in an easy and complex version, respectively. Task difficulty was validated in a previous rating study (see 2.2.1.2).

The participants were recruited on campus of Bielefeld University and received either partial course credit or chocolate bars as compensation for their participation. The study was approved by the ethics committee of Bielefeld University.

2.2.1.2 Performance tasks

Anagram solving

Anagram-solving tasks are quite common in social facilitation research (Aiello & Svec, 1993; Davidson & Henderson, 2000; Park & Catrambone, 2007). Anagrams are nonsense-words, as their letters are positioned in random and incorrect order. Thus, the task is to rearrange the letters back into the correct order (e.g. “riecpn” = “prince”). Here, participants had to solve 5-letter anagrams as quickly and accurately as possible. In order to create the two difficulty categories, a previous rating study was conducted, from which the 10 anagrams with the fastest and slowest solution times each were chosen out of a pool of 112 different anagrams. Accordingly, the two conditions differed significantly, $t(9) = 6.47, p < .001, d_z = 2.04$.

In the current study, the 20 anagrams were individually presented on the computer screen in randomized array. The participants were asked to enter the correct solution of each presented anagram via the computer keyboard. This enabled to record both reaction time (time from stimulus presentation until first keypress of the correctly solved anagram) and accuracy (ratio of correct solutions to total number of anagrams) as dependent variables.

Numerical distance

The numerical distance effect described by Moyer and Landauer (1967) was used as paradigm for this task. The effect states that the reaction time for a comparison of the numerical value of two digits is inversely proportional to the numerical distance between them. That is, the larger the numerical distance between two digits, the faster and more accurate is the respective comparison response.

In the current study, the task was to decide by means of a specified keypress whether a presented number was bigger or smaller than “5”. Regarding the construction of an easy and complex condition, for the easy version of the task, these numbers were either “1” or “9”, and for the complex version “4” or “6”. A previous rating study confirmed significant differences between these two categories, $t(7) = 4.43$, $p = .002$, $d_z = 1.57$. Both difficulty categories consisted of 10 trials each, and the resulting 20 trials total were presented in fully randomized order. Response latency served as dependent variable.

Finger tapping

For the finger-tapping task, participants were asked to tap a recurring key sequence on the computer keyboard as quickly and accurately as possible for the duration of 90 seconds. The key sequence consisted of either two (for the easy condition, “S” and “L”) or six (for the complex condition, “E”, “K”, “V”, “A”, “P”, “X”) alternating keys and was presented on the computer screen in front of the participants. Again, a previous rating study confirmed significant differences between the two conditions, $t(7) = 4.68$, $p < .001$, $d_z = 1.65$. The total number of taps served as the dependent variable. The order of the two conditions within the task was randomly assigned.

Motoric tracking

The motoric-tracking task required the participants to follow an array of square buttons, which were presented on their computer screen, using the computer mouse. Concretely, they were asked to click on each square box that appeared at a random position on the screen as quickly as possible. Every click initiated the appearance of the subsequent box at a random position on the computer screen. One trial consisted of a sequence of six consecutive box presentations. As 10 trials were run, this resulted in 20 fully randomized trials total.

Task difficulty was manipulated by asking the participants to either hold the mouse in their right hand (for the easy condition, as all participants were right-handed) or in their left hand (for the complex version) before each new trial. Additionally, the box buttons in the easy condition were four times larger than the buttons in the complex condition, which taken together led to significant differences between the two task versions, as confirmed by a previous rating study,

$t(7) = 13.20$, $p < .001$, $d_z = 4.67$. The total solution time per trial (time from onset of first box until offset of sixth box) was recorded as dependent variable.

Order of presentation

The order of the four tasks within the experiment was fixed, starting with the finger-tapping task, followed by the numerical distance task, the motoric-tracking task and lastly, the anagram-solving task. Carry-over effects were regarded as negligible, as the tasks aimed at diverse cognitive and motoric domains, the task order was standardized over the three presence groups, and the relevant analysis comparisons were between-groups and not between-tasks. LiveCode (RunRev Ltd.) software was used for task development and conduction of the experiment.

2.2.1.3 Presence conditions

The anthropomorphic robot head Flobi was used to create the robotic presence condition. The robot has technically (Lütkebohle et al., 2010) and aesthetically (Hegel, Eyssel, & Wrede, 2010) been designed at Bielefeld University. It holds 18 degrees of freedom total, from which 15 actuators alone are located in the face in order to move eyes, eyelids, brows, and lips, enabling a wide variety of specific facial displays. As a key feature, the robot head is composed of exchangeable modular parts (features like hairstyle, lips, eyebrows, etc.), which enables quick and flexible appearance changes. Because the human confederate for the human presence condition was female, a female version of the Flobi head was chosen for this study, accordingly.

Figure 1. The Flobi robot head (left, copyright: Frank Hegel, 2011) and laboratory set-up (right).

The behavior of the robot and human confederate followed a previously defined time-based behavior script to ensure comparability of the two presence conditions. The robotic behavior was pre-programmed, whereas for the human presence condition, thorough training, an additional pre-study without a robot condition, and previous behavior ratings ensured reliability of the confederate's behavior throughout the study. To rule out potential emotional or interactional confounding factors, a mere presence behavior design was employed, that is, no emotions were displayed by either the confederate or the robot, and interactional behavior was neither initiated nor responded to. The human confederate and the robot were simply quiet, non-engaging spectators.

Upon entering the laboratory, depending on the presence condition, the robot or the human confederate quickly glanced towards the participant, nodded slightly in that direction and went back to their tasks on their own computer screen, which was positioned opposite to the participant. Over the course of the experiment, they glanced towards the participant every 3-4 minutes. The rest of the time, they gazed at their own computer screen with short interruptions, when they quickly glimpsed up at the room and changed their facial expressions from a rather "neutral" to "concentrated" look back and forth several times. The robot and the female confederate were unknown to all participants.

2.2.1.4 Procedure

The experiment took place in a laboratory at Bielefeld University. The acoustically shielded chamber (approximately 2.5 m × 3.5 m) contained a single desk that the participants either sat at alone or opposite to the robot or human confederate, depending on the presence condition. Upon arrival, participants were informed that, as part of a software evaluation, they had to complete four different tasks on the computer and answer a questionnaire afterwards. Additionally, in the two presence conditions, they were further told that a technical assistant or a robot (depending on the condition) was going to be sitting in the room with them to either monitor certain software parameters on-line on their own computer screen or a specific visual task, respectively, which was explained as being a part of a cooperative study with an informatics group. The relevant criterion for the two cover stories was that both the robot as well as the human confederate had an equally engaging visual task to attend to on their own computer screen. Hence, participants were asked not to communicate with the attendant during the experiment.

Following these instructions, the participants entered the laboratory, where the experimenter started the experimental software and left the room. After completion of the last performance task, participants were led to an adjacent room for a manipulation check form, that asked how observed they felt during the experiment via a 10 cm paper-pencil visual analogue scale. Lastly, the experimenter revealed the true purpose of the study. None of the participants retrospectively reported suspicion regarding the cover story.

2.2.2 Results

As the key findings in the social facilitation paradigm can also be expressed as a large performance difference between the easy and complex task version for presence conditions (due to predicted performance enhancement on easy and performance impairment on complex tasks) compared to a relatively smaller performance difference for the alone control condition, difference scores between the easy and complex condition of every task were calculated for the three presence groups. Then, a one-factorial ANOVA for the presence factor (alone vs. human presence vs. robotic presence) was conducted to test for significant main effects of presence. Subsequent planned comparisons were computed to reveal potential between-group contrasts – alone vs. both presence groups (one-sided) and, relevant for the second hypothesis – whether human and robotic presence differed from each other, human presence vs. robotic presence (two-sided).

2.2.2.1 Anagram solving

Regarding task accuracy, the ANOVA revealed that the three presence groups significantly differed from each other, $F(2,103) = 3.61$, $p = .030$, $\eta_p^2 = .07$ (see Fig. 2 and Table 1). Planned comparisons further indicated a significant difference between the alone group compared to the two presence conditions, $t(103) = 1.91$, $p = .029$, $r = .19$. That is, a larger difference in accuracy between easy and complex anagram trials was found when a robot or human was present compared to the alone control group. The contrast between human and robotic presence was marginally significant, $t(103) = 1.92$, $p = .057$, $r = .19$. In consequence of the significant differences between the presence groups in the accuracy scores, it was refrained from analyzing response latencies, due to the mere fact of missing reaction time data, especially in the complex condition.

2.2.2.2 Numerical distance

Concerning response latencies in the numerical distance task, median reaction times varied significantly with the presence condition, $F(2,103) = 3.96$, $p = .022$, $\eta_p^2 = .07$ (see Fig. 2 and Table 1). Again, the two presence groups revealed a larger difference score than the alone control group, $t(103) = 2.47$, $p = .007$, $r = .24$. Human and robotic presence however did not produce significantly different performance effects, $t(103) = 1.30$, $p = .196$, $r = .13$.

2.2.2.3 Finger tapping

Regarding the finger-tapping task, the ANOVA did not uncover a significant main effect of presence on the total number of taps, $F(2,103) = 2.19$, $p = .117$, $\eta_p^2 = .04$, nonetheless, planned comparisons indicated the same pattern of results as for the previous tasks (see Fig. 2 and Table 1). The difference in the total amount of taps was significantly larger between easy and complex version of the task for the presence conditions compared to the alone group, $t(103) = 1.88$, $p = .031$, $r = .18$, whereas robotic and human presence again did not significantly differ from each other, $t(103) = 0.96$, $p = .340$, $r = .09$.

2.2.2.4 Motoric tracking

No significant results could be obtained for the motoric-tracking task. Neither ANOVA, $F(2,103) = 0.88$, $p = .416$, $\eta_p^2 = .02$, nor planned comparisons (alone vs. presence: $t(103) = 0.57$, $p = .285$, $r = .06$; human vs. robotic presence: $t(103) = 1.19$, $p = .236$, $r = .12$) could uncover any significant effect of presence on the performance differences between the easy and complex task version. Despite testing well in the previous validation study (see 2.2.1.2), these null results are probably due to large within-group variances that could stem from difficulties in task conduction during the course of the experiment, misunderstanding of task instructions, or even cheating in order to finish the task faster. Thus, this task will be excluded from further interpretation and discussion of the results.

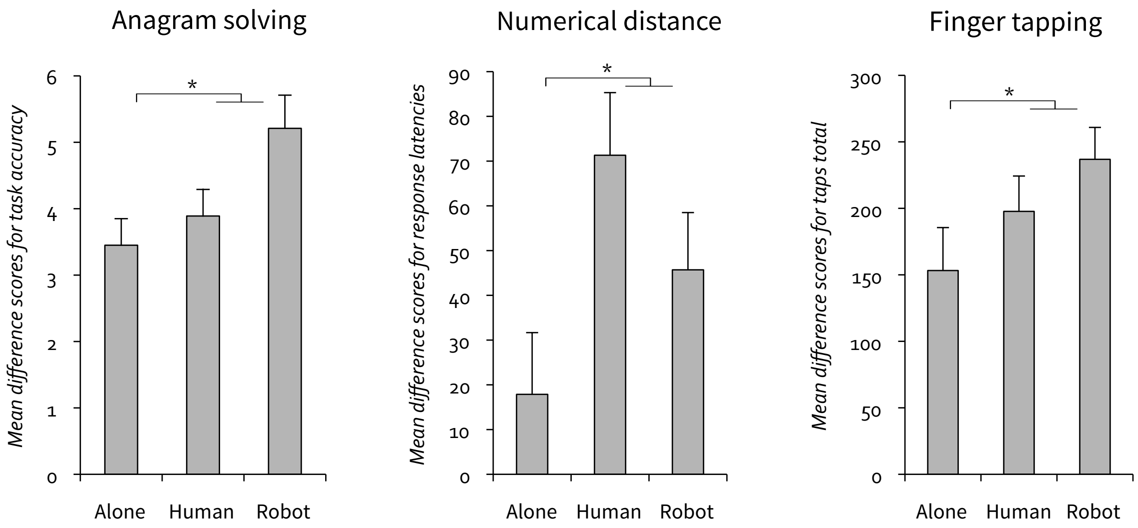

Figure 2. Mean difference scores between easy and complex task version of anagram-solving, numerical distance, and finger-tapping task. Error bars represent standard error of mean.

2.2.2.5 Subjective observation impression

Concerning the manipulation check regarding the subjective perception of having been monitored during the experiment, significant overall effects of presence were discovered, $F(2,103) = 6.35$, $p = .003$, $\eta_p^2 = .11$. These matched the previous performance results in terms of a higher subjective observation impression during the two presence conditions compared to the alone control condition, $t(103) = 2.82$, $p = .003$, $r = .27$. Interestingly, the planned comparisons also revealed a significant difference between the presence of the human confederate versus the presence of the robot, *t*(103) = 2.23, *p* = .028, *r* = .21, with robotic presence leading to higher observation impression than human presence. Participants thus felt most observed when a robot was present (*M* = 5.24, *SD* = 3.42), followed by human presence (*M* = 3.51, *SD* = 2.81) and the alone condition (*M* = 2.56, *SD* = 3.33).

Table 1: Descriptive performance results for anagram-solving task, numerical distance task, and finger-tapping task.

	Anagram solving (M and *SD)*[a]		*Numerical distance (Md* and *SD)*[b]		*Finger tapping (M* and *SD)*[c]	
	Easy	Complex	Easy	Complex	Easy	Complex
Alone	9.71 (0.57)	6.26 (2.91)	561.21 (166.22)	579.08 (163.00)	325.74 (189.57)	172.50 (73.54)
Human	9.80 (0.41)	5.91 (2.75)	527.80 (89.62)	599.11 (142.41)	341.46 (154.59)	143.74 (48.68)
Robot	9.55 (0.71)	4.33 (3.21)	534.88 (105.81)	580.58 (119.99)	370.91 (154.31)	134.06 (40.95)

[a] Depicted are number of correctly solved anagrams.
[b] Depicted are response latencies in milliseconds.
[c] Depicted is the total amount of taps.

2.2.3 Discussion

This first study investigated whether social influence of non-human entities could be measured indirectly using human performance as objective quantification. This was of special interest as the social influence would have to be very profound in order to produce the predicted effects of the social facilitation paradigm in a mere presence (non-interactional) setting. Hence, it was of further interest how comparable potential effects of robotic presence were to human presence. Therefore, three groups (human confederate present, robot present, alone control condition) were tested on different cognitive and motoric tasks with manipulated task difficulty (easy, complex) in order to investigate whether the social facilitation effects could be replicated for robotic presence.

Overall, the present results replicate the social facilitation effects both for human and robotic presence. That is, the performance differences between easy and complex task versions were significantly greater when a human or robot was present compared to the alone control condition, indicating performance enhancement on easy tasks as well as performance impairment on complex tasks for mere robotic and human presence. Interestingly, human and robotic presence did not significantly differ from each other on the performance measures. The mere presence of a social robot indeed seems to lead to similar performance effects as human presence, which emphasizes the profoundness of its social influence.

Furthermore, regarding the subjective impression of having been monitored, the two presence conditions consequently differed significantly from the alone control group – participants felt more observed when a robot or human was present compared to when they were alone. However, a significant difference additionally emerged between human and robotic presence, the latter eliciting the strongest observation impression. Taken together, the subjective results match the performance measures in terms of an apparent perception of co-presence in both presence conditions, which is crucial for the emergence of social facilitation effects. The strongest observation impression obtained for the robot condition however is striking, as this does apparently not similarly extend to stronger performance effects of robotic versus human presence.

How can the difference in the subjective measure be explained? The higher observation scores for robotic presence compared to human presence might be due to surveillance effects attributable to the two cameras in the robot's eyes. This might have increased situational awareness or even left room for the assumption that an additional person might also be watching through the robot's eyes, following the rationale that social responses towards non-human agents might just be observed responses directed at the person behind the interface. This study cannot address this issue directly, however, examining whether the obtained effects remain for a presence condition in which the robot is switched off or in which its eyes are closed, thus disabling any surveillance assumptions, could approach this assumption.

Another explanation might be novelty effects. None of the participants in the robot group reported previous interaction with a robot or knowledge of the Flobi robot head. Based on this unfamiliarity, the robotic presence might have been more distracting, but also more salient than the rather familiar presence of another human. To address this factor, it appears enlightening for future research to include a condition with additional familiarization with the robot, for example through a previous interaction. If the current findings could simply be reduced to a novelty effect of robotic presence, then future findings should either find no significant differences between human and robotic presence with previous familiarization or the results for the familiarization condition should lie below those of the human presence condition, but not above them, as is the case here for the subjective observation impression. This approach, on the other hand, might counteract the subtlety of the current unobtrusive mere presence study design, potentially leading to overall heightened situational awareness and consequential incalculable effects. Other measures thus might appear more suitable for tackling familiarization issues.

A further relevant aspect for the current findings could be presumably higher unpredictability of the robot compared to the human confederate. As participants did not have any prior knowledge of the robot‘s potential behavior range or of how sophisticated its "cognitions" might be, they likely did not possess an elaborate mental model of the robot and had only the robot's observables to infer any functionality ascriptions (Duffy, 2003). Thus, the robot's presence could possibly have heightened uncertainty, which should result in enhanced arousal, heightened alertness, and monitoring-responses (see Lynn, 1966; Norman, 1980; Zajonc, 1980). Recent studies further suggest unpredictability as mediator of anthropomorphism in human-robot interaction (Epley, Akalis, et al., 2008; Waytz, Morewedge, et al., 2010), with an unpredictable robot leading to higher anthropomorphization, which links nicely to the current findings. Altogether, unpredictability seems a potent influence factor for the emergence of social responses towards nonhuman entities in general, as the research on anthropomorphization points out, and also for the obtained results of the current study, considering the effects of unpredictability itself on arousal. However, the strongest observation impression obtained for the robot condition does apparently not result in equally stronger performance effects of robotic versus human presence.

How would social facilitation theory explain the current results? Considering a unified, interactive perspective of the several explanatory approaches, according to drive aspects, robotic and human presence equally elevated participants' arousal, which caused the current performance differences compared to the alone condition. Cognitive process approaches emphasize distraction and attentional conflicts, which seem likely influence factors for the current findings. It has to be noted, however, that the performance differences largely stem from a performance decline on the complex task versions to a relatively smaller performance increase on the easy task versions (probably due to common ceiling effects in social facilitation research). This, in turn, renders these cognitive factors even more substantial, as shared attention as well as distraction appear especially detrimental concerning performance on complex tasks.

Social comparison approaches would make evaluation apprehension and impression management mechanisms elicited by the human confederate as well as the robot accountable for the effects. This means that participants should have experienced apprehension about how the robot and the human would evaluate them and wanted to present themselves in a socially acceptable manner in both presence groups. Of course, this in turn might again lead to heightened arousal, distraction, and other attentional conflicts. But would we really try to impress a robot? Additionally, as already mentioned, it appears quite interesting that even though the observation impression scores for robotic presence were significantly higher than those for human presence, these variations did not equally apply to the performance results, as both presence groups did not differ from each other on these measures. The confounding of this explanatory mechanism with the other approaches seems unsolvable for human presence. However, now that the performance data prove a comparable social influence of a humanlike robot to that of a real human, it appears possible to disentangle the explanatory overfit of the models behind the social facilitation effect. A switched-off robot for example should diminish evaluation apprehension, whereas arousal or cognitive distraction mechanisms may still prevail.

Linking this intent to the exploration of the unpredictability account and the potential surveillance factor, it yet again appears sensible to integrate a condition in the paradigm in which the robot is obviously switched off, thus ruling out influences due to unpredictable robotic behavior, decreasing participants' uncertainty regarding the robot to a minimum as well as minimizing social comparison and evaluation apprehension mechanisms. This, additionally, appears especially relevant for applied settings, as one implication drawn from the current findings is that social facilitation effects need to be taken into account not only for the development and application of robotic platforms in workplace, tutoring, or other assistance scenarios, but also for further research on human-robot interaction in general, as soon as task performance is of interest. Even mere robotic presence, as well as mere human presence, may already impair human performance, specifically on new or complex tasks. Hence, omnipresent robotic systems do not seem necessarily advisable and may even induce counterproductive effects. Conversely, the employment of robotic systems may have motivating effects just by being present when an easy or well-learned task has to be performed. For the application of robots in scenarios in which human performance is of essential value, every decision of presenting a robotic system should therefore include thoughtful considerations concerning the respective task difficulty. One option that would be highly useful for the applied setting therefore seems to be the possibility to switch off the robot and thereby eliminating its social influence when necessary, thus ruling out any potential effects on human performance.

Connecting this rationale to the aforementioned possibilities to investigate the influence of unpredictability, surveillance effects, and social comparison explanations, consequently, a second study was conducted that contained a condition in which the robot was obviously switched off.

2.3 Study 2 – Effects of Animacy on the Social Facilitation Paradigm with Social Robots

2.3.1 Methods

2.3.1.1 Participants and design

To investigate whether the social facilitation effects obtained in Study 1 would also occur in the presence of a robot that was switched off, a 3 (*presence type:* alone, active robot present, switched-off robot present) × 2 (*task difficulty:* easy, complex) design analogous to that of Study 1 was realized. Sixty-four participants (38 women, no gender differences between groups) with a mean age of 23.48 years (*SD* = 3.67 years) were recruited on campus of Bielefeld University and randomly assigned to one of the three groups. As in Study 1, the performance on four different computerized tasks (see 2.3.1.2), which were administered in both an easy and complex condition served as dependent variable. Participants received either partial course credit or monetary compensation in return for their participation.

2.3.1.2 Performance tasks

Except for the motoric-tracking task, the same performance tasks as in the previous study (see 2.2.1.2 for detailed descriptions) were utilized for the current experiment. Thus, participants had to solve easy and complex 5-letter anagrams as quickly and accurately as possible with task accuracy being the dependent variable (anagram solving). Further, in the numerical distance task participants were asked to quickly respond whether a presented number was bigger or smaller than “5” both in an easy and complex version. Response latency served as dependent measure. Regarding the finger-tapping task, participants were required to tap an easy and complex recurring key sequence as quickly and accurately as possible for 90 seconds with the total amount of taps serving as dependent variable. Despite promising results in a previous rating study, the motoric-tracking task did not reveal any differentiating effects in Study 1 (see 2.2.2). This might have been due to difficulties with the task instructions or purposeful disobedience due to the more laborious task requirements compared to the other tasks. Therefore, a replacement task was developed for the current study.

Visual search task

The visual search task was based on the classical work by Treisman and Gelade (1980) on visual search processes, specifically aiming at feature- versus conjunction-search as operationalization of the easy and complex condition. During feature-search, the detection of one target feature, e.g. shape, color, orientation, etc. is sufficient to quickly pick out a specified target in an array of distractors not sharing this specific property, for instance detecting a blue cross in an array of all yellow distractors, which leads to "pop out", bottom-up detection of the target. However, when visual features of target and distractor stimuli overlap (e.g. detecting a blue cross in an array of yellow crosses and blue circles), the target becomes harder to identify and requires top-down search processes, resulting in slower detection times, which is correlated to the number of presented distractors.

Hence, for the current task, participants had to quickly detect a specified target letter in an array of 45 randomly scattered stimulus letters across the screen. For the easy condition, the target stimulus contained a pop-out property, which was balanced across the easy trials, that is, the target was either the blue letter "T" or "X" or the brown or green letter "S" among 44 distractor stimuli, of which half were the brown letter "T" and the other half the green letter "X". For the complex condition, the target stimulus was a green "T", thus sharing visual features with both the distractor stimulus letters. The task contained 10 trials per difficulty category; furthermore, six trials were included that did not contain a target stimulus in order to promote both fast but also accurate responses. As soon as they detected the target, participants had to press a previously defined key with their dominant hand, otherwise, they had to respond with another keypress in case the target was absent or could not be found. The stimulus arrays for every trial were previously defined and presented in complete random order. A previous rating study confirmed significant differences between the two difficulty conditions, $t(9) = 8.49$, $p < .001$, $d_z = 2.68$. Response latencies served as dependent measure.

Order of presentation

As before, the task order within the experiment was fixed, starting with the finger-tapping task, followed by the numerical distance task, visual search task, and finally the anagram-solving task. Again, LiveCode (RunRev Ltd.) software was used for the implementation and conduction of the experiment.

2.3.1.3 Presence conditions

As in the previous study, the Flobi robot head was used for the two robotic presence conditions. In order to keep the robot equally visible across the two presence conditions and to enable a clearer distinction between the two conditions, it was refrained from placing the robot behind a computer screen. For this study, the robot head was positioned directly on the table opposite of the participant, so that both robots were clearly visible throughout the whole experiment.

Figure 3. The adjusted Flobi robot head in the switched-off condition.

The switched-off robot's head was leaning to one side (see Fig. 3), both eyelids half-closed, but not on the same level, however pulled down enough so that the cameras in the robot's eyes were not visible anymore. Furthermore, the eyebrows were turned askew and the plug of the power cord was positioned visibly on the table, resulting in an overall impression of an inactive, inanimate, turned-off robot. A previous rating study confirmed that, in this setup, the switched-off robot was indeed accurately perceived as off ($N = 10$, $\chi^2 = 10.00$, $df = 1$, $p = .002$).

The active robot on the other hand performed a previously implemented time-based behavior script comparable to that of Study 1. However, as there existed no computer screen to focus the robot's gaze on, the robot looked around the room the whole duration of the experiment with short glances towards the participant every 3-4 minutes. Its facial expression changed from a rather "neutral" to a "concentrated" look several times.

2.3.1.4 Procedure

The procedure of the experiment was similar to that of Study 1 (see 2.2.1.4). Participants were asked to complete four different computerized tasks, supposedly as part of a software evaluation study. In the two presence conditions, they were further told that a robot was going to be sitting in the room with them, as the laboratory was shared with an informatics group, however, that this should not bother them for this experiment. Additionally, participants in the group with the switched-off robot were explicitly told that the robot was turned off. In case of the active robot, however, participants were informed that the robot was running in a long-term pilot phase and thus would be active, however, that they should not let this distract them. In order to not interfere with the aim of the robotic test phase and that of the current software evaluation, they were asked not to engage with the robot.

After these instructions, the experimenter started the experimental software and left the room. Upon completion of the last task, participants filled in the subjective observation impression form in an adjacent room and were debriefed by the experimenter. None of the participants reported suspicion regarding the cover stories.

2.3.2 Results

As the descriptive data did not clearly replicate the predicted pattern of performance enhancements on the easy task versions and performance impairments on the complex versions for the active robot group compared to the alone condition similar to Study 1, it was refrained from calculating difference scores between the easy and complex tasks versions. Instead, one-factorial ANOVAs for the presence factor (alone vs. switched-off robot present vs. active robot present) were separately calculated for the two difficulty conditions, followed by planned comparisons to reveal potential between-group contrasts – alone vs. both presence groups (one-sided) and switched-off robotic presence vs. active robotic presence (two-sided).

2.3.2.1 Anagram solving

Regarding task accuracy, the three presence groups turned out to significantly differ from each other on the easy task version, $F(2,59) = 3.72$, $p = .030$, $\eta_p^2 = .11$ (see Fig. 4 and Table 2). Subsequent analyses revealed a significant contrast between the alone control condition and the two robotic presence groups, $t(59) = 3.32$, $p = .001$, $r = .40$, whereas the effects induced by the active robot did not significantly differ from those of the switched-off robot, $t(59) = 1.12$, $p = .270$, $r = .14$. Concerning the complex task version, no significant differences between the groups were obtained, $F(2,59) = 0.52$, $p = .597$, $\eta_p^2 = .02$, likely caused by floor effects that may have manifested due to the sample size.

However, as noted above, regarding the descriptive statistics, the apparent differences on the easy task version run in the opposite way of the predicted social facilitation effects – here, both presence groups induced a descriptive impairment of task accuracy compared to the alone control condition, comparable to the (yet non-significant) pattern found for the complex task version.

Table 2: Descriptive performance results for anagram-solving task and numerical distance task.

	Anagram solving (M and *SD)*[a]		*Numerical distance (Md* and *SD)*[b]	
	Easy	Complex	Easy	Complex
Alone	9.95 (0.23)	5.63 (2.38)	523.10 (60.04)	562.20 (74.84)
Robot off	9.42 (0.77)	4.74 (2.92)	596.45 (88.28)	636.05 (99.85)
Robot active	9.67 (0.64)	5.29 (2.84)	590.48 (92.46)	647.83 (115.41)

[a] Depicted is the amount of correctly solved anagrams.

[b] Depicted are response latencies in milliseconds.

2.3.2.2 Numerical distance

Median reaction times differed significantly between the presence conditions, both on the easy, $F(2,60) = 5.01$, $p = .010$, $\eta_p^2 = .14$, and complex version, $F(2,60) = 4.54$, $p = .015$, $\eta_p^2 = .13$, of the task (see Fig. 4 and Table 2). Planned comparisons further indicated significant contrasts between the two presence groups compared to the alone control group, again, both for the easy task version, $t(60) = 3.16$, $p = .001$, $r = .38$, as well as the complex task version, $t(60) = 2.97$, $p = .002$, $r = .36$. Yet, the two effect patterns appear similar, that is, response latencies were significantly larger on both difficulty versions for the two robotic presence groups compared to the alone condition. Further, no significant contrasts were found between the two robot groups (easy task version: $t(60) = 0.24$, $p = .813$, $r = .03$; complex task version: $t(60) = 0.39$, $p = .699$, $r = .05$). Both robotic presence conditions thus produced comparable performance effects.

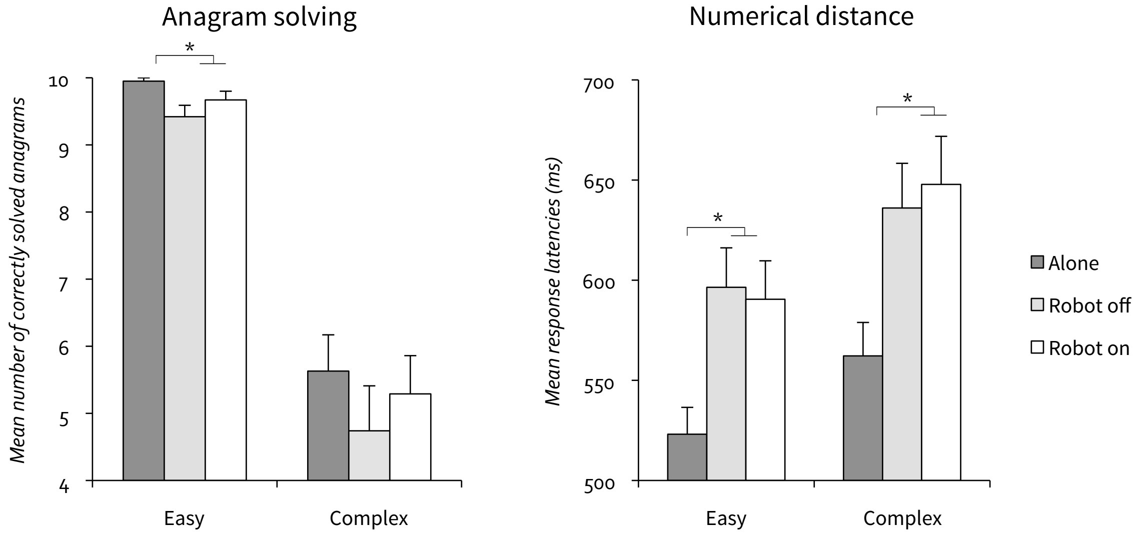

Figure 4. Performance results for easy and complex task version of the anagram-solving and numerical distance task. Error bars depict standard error of mean.

2.3.2.3 Finger tapping

Regarding the total amount of taps, the three presence groups marginally differed on both difficulty versions of the task (easy task version: $F(2,59) = 2.61$, $p = .082$, $\eta_p^2 = .08$; complex task version: $F(2,59) = 3.16$, $p = .050$, $\eta_p^2 = .10$), see Fig. 5 and Table 3. However, planned comparisons showed that the number of taps total was significantly larger for the alone control condition compared to the two presence groups, both for the easy, $t(59) = 2.04$, $p = .023$, $r = .26$, and complex, $t(59) = 2.50$, $p = .007$, $r = .31$, difficulty condition.

Again, the effects induced by the switched-off robot did not significantly differ from those of the active robot, as subsequent contrast tests revealed (easy task version: $t(59) = 1.15$, $p = .256$, $r = .15$; complex task version: $t(59) = 0.43$, $p = .666$, $r = .06$).

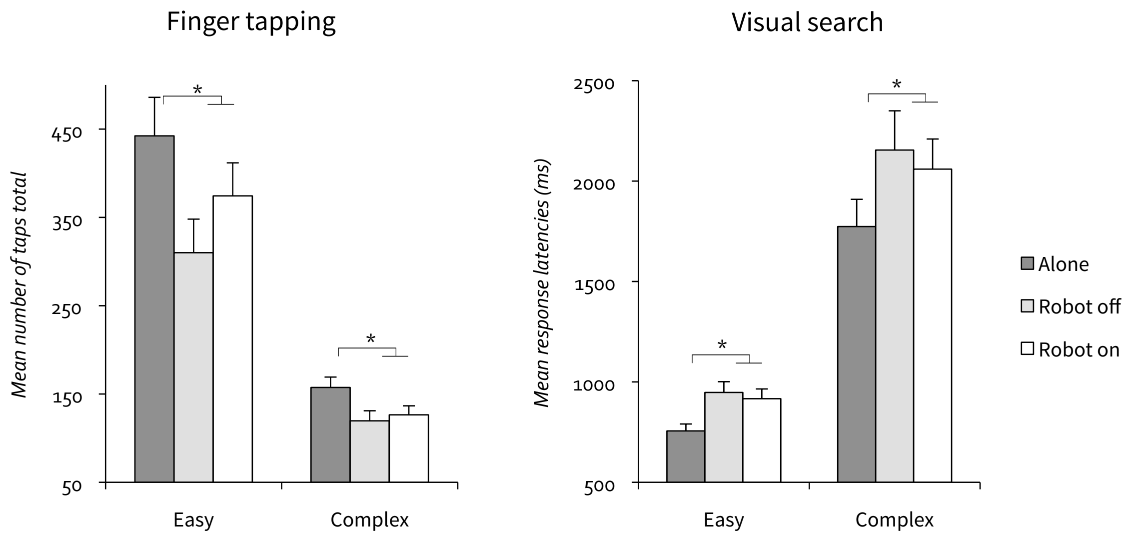

Figure 5. Performance results for easy and complex task version of the finger-tapping and visual search task. Error bars represent standard error of mean.

2.3.2.4 Visual search

The response latencies on the visual search task differed significantly for the easy task version, $F(2,59) = 4.65, p = .013, \eta_p^2 = .14$, yet again with the two presence groups producing significantly larger response latencies than the alone control condition, $t(59) = 3.02, p = .002, r = .37$ (see Fig. 5 and Table 3). The active robot did not produce significantly different performance variations than the switched-off robot, $t(59) = 0.47, p = .641, r = .06$.

Regarding the complex task version, no significant results could be obtained regarding the overall ANOVA, $F(2,59) = 1.46, p = .240, \eta_p^2 = .05$. However, the two robotic presence groups differed from the alone control condition on the planned contrasts, $t(59) = 1.67, p = .049, r = .21$, leading to designated performance decreases compared to the alone group, whereas not significantly differing from each other, $t(59) = 0.42, p = .677, r = .05$.

Table 3: Descriptive performance results for the finger-tapping and visual search task.

	Finger tapping (M and *SD)*[a]		*Visual search (Md* and *SD)*[b]	
	Easy	Complex	Easy	Complex
Alone	442.35 (195.37)	157.40 (53.06)	755.95 (154.67)	1773.60 (606.47)
Robot off	310.05 (165.66)	119.63 (49.71)	947.40 (240.49)	2154.60 (874.12)
Robot active	374.39 (179.78)	126.43 (49.13)	916.36 (233.95)	2059.59 (703.94)

[a] Depicted is total amount of taps.
[b] Depicted are response latencies in milliseconds.

2.3.2.5 Subjective observation impression

The subjective perception of having been monitored during the experiment differed significantly between the three groups, $F(2,61) = 3.79, p = .028, \eta_p^2 = .11$. Expectably, the two robotic presence conditions elicited a significantly greater observation impression than the alone control group ($M = 1.89$, $SD = 2.79$), $t(61) = 1.90$, $p = .031$ (one-sided), $r = .24$, matching the performance results. Additionally, the switched-off robot ($M = 2.52$, $SD = 2.74$) differed significantly from the active robot ($M = 4.14$, $SD = 2.89$), which reasonably evoked the highest scores on this measure, $t(61) = 1.90$, $p = .031$ (one-sided), $r = .24$. However, a subsequent comparison revealed that, in fact, participants in the switched-off robot group did not significantly differ from those in the alone control group concerning their impression of having been monitored, $t(61) = 0.71$, $p = .240$ (one-sided), $r = .01$.

2.3.3 Discussion

This second study examined how far the effects of robotic presence on human performance were reaching regarding the comparable influence of mere robotic presence to mere human presence on performance that Study 1 documented. Would the presence of a switched-off robot also induce social facilitation effects? Hence, three groups (switched-off robot present, active robot present, alone control condition) were tested on four different cognitive and motoric computerized tasks that were manipulated in terms of difficulty (easy, complex).

Overall, the present findings indicate a significant effect of robotic presence on task performance both for the active as well as the switched-off robot. Compared to the alone control condition, participants in the two robot groups obtained significantly lower results on the performance tasks; however, this held true for the easy as well as the difficult task versions. This effect pattern might be caused by the smaller sample size than that of Study 1, allowing more influence of inter-subject variability, but also floor and ceiling effects, which are common methodological problems in social facilitation research (Bond & Titus, 1983). Another explanation for this result pattern might be that the current study setup potentially induced more pronounced social presence as the robots were positioned directly opposite to the participant without the barrier of their "own" computer screen. Especially for this more obvious or salient account on presence manipulation compared to the very conservative setup of Study 1, the easy conditions of the administered tasks might actually not have been easy enough for the classical social facilitation effects to occur as they were still novel tasks that were not overlearned yet, despite them greatly differentiating from the respective complex versions in the pre-tests (see 2.2.1.2 and 2.3.1.2).

Taken together, the current study produced clear evidence of performance decreases, or social inhibition effects, when participants had to solve cognitive and motoric tasks in the presence of both an active robot and, similarly, one that was obviously switched off. Additionally, the two robotic presence conditions apparently did not induce qualitatively different effects on human performance, as the results on the performance tasks did not significantly differ from each other.

However, regarding the subjective impression of having been monitored, participants indicated a significantly higher observation impression for the active robot condition compared to the switched-off robot condition. Further, participants in the switched-off robot condition did not feel significantly more observed than participants in the alone control group. This might seem trivial concerning the fact that the robot's eyelids in the switched-off condition were indeed lowered so much that the cameras in the robot's eyes were not visible any more and thus, there was simply no chance that the robot, despite being inactive, could have monitored the participants. Yet, this finding is of special interest because participants showed precisely the same amount of performance impairments on the administered tasks in both robotic presence groups, despite the fact that they felt significantly less monitored when the robot was switched off. This allows for two tentative conclusions: Firstly, effects of social presence on performance might work along a threshold of influence, that is, above the threshold, various forms of presence might have a uniform impact, thus not differentiating along the performance dimension. Secondly, this threshold of influence seems to lie very low, as even a switched-off robot apparently induces enough social presence to produce performance variations.

This connects nicely to a classical study by Schmitt, Gilovich, Goore, and Joseph (1986), who demonstrated that even the presence of a blindfolded person wearing earphones was sufficient to induce social facilitation effects. Apparently, the profound mechanisms underlying these responses are transferrable to robotic presence as well.

Hence, these current findings emphasize the explanatory power of potential arousal increases and attentional constraints on the social facilitation effects, as the influence of evaluative mechanisms was clearly, although potentially not completely, diminished, echoing the early propositions made by Zajonc (1980, see above) on the profoundness of the mere presence design.

As already discussed in relation to the results of Study 1, unpredictability might be a suitable source of heightened arousal but also distraction or attentional conflicts. This seems to be an obvious influence factor regarding the active robot, of which participants likely did not have formed an elaborate mental model yet, as its active behavior should have led to heightened alertness and also monitoring responses (Lynn, 1966; Norman, 1980; Zajonc, 1980). These in turn most likely entail deteriorative effects on cognitive performance, as they provoke attention shifts and distraction. In this light, the consistent performance impairments on the complex as well as the easy task versions appear quite sensible, regarding the fact that participants might have been universally distracted from fulfilling the task demands due to the salient robotic presence. This consideration might also hold true for the switched-off robot, as based on unfamiliarity assumptions, even the switched-off robot might have induced uncertainty, resulting in heightened alertness, monitoring-responses, and distraction. For example, participants might have been unsure whether the robot could have been able to "wake up" by itself again. Or else, related to the unfamiliarity line of reasoning, curiosity might have led them to observe the robot more closely, especially when they did not feel that watched by it in the switched-off condition, again resulting in distraction and attention shifts.

Despite these suppositions, the findings of the previous studies clearly indicate that the mere presence of a humanlike entity, active or not, truly human or not, can already have drastic effects on human performance. Additionally, a simple surveillance factor due to potential visibility of cameras, which was discussed in relation to the findings of Study 1, as well as responses directed at the human behind the machine, can most likely be ruled out as (primary) explanatory mechanisms. The latter alludes to one specific issue regarding social responses towards non-human agents, namely the alleged question at whom exactly these social reactions are directed – what if social reactions towards artificial humanlike entities only occur because they are directed at the human mind behind the non-human agent? Prior research tried to shed light onto this question and although results have been mixed, these studies suggest that the factor of agency cannot solely be attributed to cause differential social responses (Blascovich et al., 2002; Nowak & Biocca, 2003). Sundar and Nass (2000) might have presented some reconciliation by showing that participants reacted indeed differently towards the computer compared to the programmer behind it, however they did not react inevitably *less social*. The current findings fit well into this line of thought.

Apart from this, the results of the current investigation demonstrate that even with a mere presence design, a more salient form of presence (here: positioning the robot directly in front of the participant without another computer screen as it was the case in Study 1) might already hamper the performance enhancements on easy tasks in favor of overall performance decreases, independent of task difficulty, due to the profound influences of cognitive and arousal factors on these effects.

This certainly appears relevant for applied settings, as the robot's influence clearly cannot simply be turned off by pulling the plug. Additionally, the effects of robotic presence might be more detrimental than the findings of the first study would suggest. The current investigation consistently showed performance impairments during robotic presence, even on supposedly easy tasks. Thus, mobile robotic platforms that could either be sent away or recognize when they need to withdraw appear quite beneficial compared to the stationary omnipresence of other potential robotic assistance systems. Taken together, it seems highly advisable to thoroughly consider whether the application of a robotic assistance system or related technology is truly necessary when human performance is of indispensable importance as long as other factors are not accounted for by subsequent research.

Relevant aspects in this regard could be the aforementioned novelty or unfamiliarity effects, as they are prone to affect perceived unpredictability and related factors like arousal enhancement and cognitive distraction, but also temporal mechanisms regarding the length of presence or repeated presence. Temporal aspects could further help to partly disentangle the overfit of explanatory constructs behind social facilitation theory. That is, drive explanation would suggest a decline of social facilitation effects over time, whereas social comparison approaches could predict a gradual increase of the effects due to potentially enhanced self-monitoring. Based on the current findings, one might argue for the former.

However, at the same time, the social facilitation paradigm does not appear fully methodologically suitable for approaching these further issues. Regarding the investigation of length of presence, time-related performance effects like a general decrease in attention and motivational decline pose a serious difficulty and potential confounding factor. Similarly, repeated presence would entail repetition or practice effects on performance. Concerning novelty effects or effects due to unfamiliarity, familiarization could also lead to a distraction of attention during the experiment and additionally foster situational awareness, thus obstructing the advantages of unobtrusive and indirect measurement of social influence. These issues possibly also account for the fact that temporal dimensions have not yet been examined thoroughly for human presence in this paradigm, which poses the additional difficulty of adequately comparing results regarding the temporal domain of human and robotic presence. A potential methodological solution could lie in the conduction of a familiarization period isolated, and ostensibly independent, from the actual experiment, yet this would of course reasonably increase the executive effort of the study.

Taken together, the unique opportunity of unobtrusive performance-based measurement of social influence that the social facilitation paradigm offers allowed for a demonstration of the apparently particularly profound nature of social responses towards non-human entities. While already ruling out several potential confounds and alternative explanatory approaches regarding the mechanisms behind social reactions towards non-human agents, this paradigm seems not entirely suitable to tackle all potential explanatory mechanisms and boundary conditions behind these responses. Yet, based on the current results, subsequent research might explore novel, elaborate design possibilities for the social facilitation paradigm.

2.3.4 Conclusion

The previous two studies aimed at taking a first step at sounding out the boundary conditions of the alleged profoundness of social responses towards non-human entities. The main conclusion that can be drawn from the current findings is that the social influence exerted by non-human entities and the social reactions towards them appear indeed extremely basal. This is underlined specifically by the fact of the employment of indirect and unobtrusive measurements in the previous experiments. Extending the scope of previous research, social influence was assessed via performance-based measures and manifested significantly compared to the control condition.

Taken together, these first two studies allow for the interim conclusions that social reactions towards non-human entities do not seem to be just overt socially learned and situationally acceptable responses, neither based purely on demand characteristics nor being reactions actually directed at the human behind the machine, as the responses in these experiments were not measured interactionally but indirectly. Participants were neither aware that they reacted socially towards a robot nor that the robot even played a role for the experiment. Instead, social responses towards non-human entities rather present extremely low-level processes, which is supported by the fact that the prime explanatory factors for the effects measured in the previous two studies lie in the arousal and cognitive process domain and do not have to be primarily tied to evaluative mechanisms or conscious processes at all. This speaks in favor of the mindless social response approach by Nass and Moon (2000), who claim that very basal cues are sufficient at accessing low-level unconscious response schemes.

The current findings already allude to a very basal, early onset and sphere of influence; however, the specifics still remain largely unclear. Additionally, the previous studies underscore the necessity to further enlarge upon the aspects of novelty and unfamiliarity, as these could have potentially confounded with the discussed unpredictability considerations. As previously mentioned, other methodological approaches than the current social facilitation paradigm seem more fruitful in subsequently tackling how basal these social responses are and what part unfamiliarity and humanlike appearance play.

These will be addressed in the following chapters. As the neural level builds the basis for all subsequent social and emotional responses, the second paradigm of this work, which is facial mimicry, will enable to particularly focus on the minimum requirements of social reactions towards non-human entities. This paradigm allows the exploration of specific appearance factors as well as variables already mentioned above, concretely, familiarity on the microlevel. At the same time, these investigations offer further insight into the functionality behind instant facial mimicry reactions regarding the debate on how supposedly unspecific and automatic these responses are versus them rather acting as a purposeful communication tool, particularly sensitive to human faces.

3. Facial Mimicry towards Robotic Faces

3.1 Theoretical and Empirical Background

Humans are social animals and the mere act of perceiving another person can have immediate effects on our own behavior. One of these direct reactions is the mimicking of a variety of the other one's observables, e.g. speech, facial expressions, or physical mannerism (Hatfield, Rapson, & Le, 2011). The presentation of pictures displaying positive or negative facial expression for example evokes corresponding facial emotional expressions in the observer. These swift and subtle mimic reactions are not necessarily visible to the eye (Lundqvist, 1995), but can be measured via electromyography (EMG) of facial muscle activity. The activation of corresponding facial muscles, termed facial mimicry, occurs spontaneously and fast (starting 300-400 ms after stimulus onset). An activation of *m. zygomaticus major* (see Fig. 6), which pulls the corners of the mouth up and back to produce a smile, can be observed in response to positive facial expressions, e.g. happy faces. Complementary, *m. corrugator supercilii* (see Fig. 6), which knits the brow into a frown, is activated during presentation of pictures displaying negative facial affect, especially sad or angry faces.

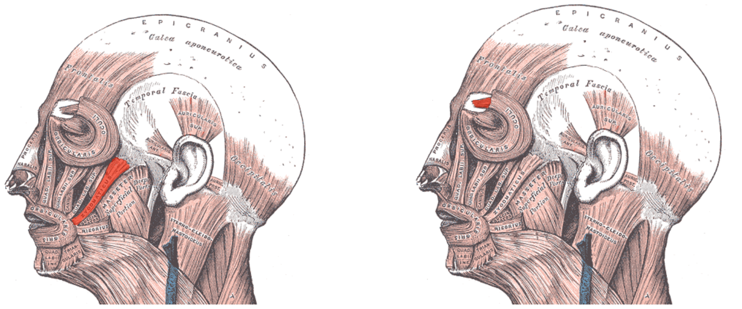

Figure 6. Location of facial muscle sites of *m. zygomaticus major* (left) and *m. corrugator supercilii* (right). Graphics are adapted from Gray (1918).

While it can be safe to assume that a prime function of facial mimicry lies in the facilitation of social interaction and interpersonal bonding, there is, however, still some ambiguity as to the nature of facial mimicry and its underlying psychological processes (Hess, Philippot, & Blairy, 1998; Moody & McIntosh, 2006; Moody, McIntosh, Mann, & Weisser, 2007).

Facial mimicry responses are important and implicit aspects of many social and emotional processes. Hatfield, Cacioppo, and Rapson (1994) regard the corresponding facial reactions as the first level of emotional contagion and empathy. According to McIntosh (1996), facial mimicry poses an antecedent of emotion through facial feedback, thus generating emotional contagion. That is, mimicking of a smile may induce the same feelings of content or happiness that are experienced by the mimicked counterpart.

As proposed by embodiment theory, cognitions are influenced by physical references, which may include facial mimicry responses (Barsalou, 1999; Thompson & Varela, 2001). As a result of activated mirror neuron processes, people may be able to use the implicit knowledge of their own physical experience to track and predict the other one's actions, which should facilitate communication with our environment, argue Wilson and Knoblich (2005).

A direct connection between empathy and facial mimicry is illustrated by Sonnby-Borgström (2002) who found that participants with high levels of empathy showed significantly more facial mimicry than low-empathic participants. Additionally, individuals with autism-spectrum disorders oftentimes fail to display fast corresponding facial mimicry reactions (McIntosh, Reichmann-Decker, Winkielman, & Wilbarger, 2006).

The fact that infants already mimic other people's facial expressions (Meltzoff & Prinz, 2002) speaks in favor of an unspecific, non-affective, automatic, and non-conscious account on facial mimicry responses. This goes in line with Chartrand and Bargh (1999), who propose an automatic direct perception-behavior link as causal chain for facial mimicry. Other researchers also suggest a direct neural link that bypasses emotional structures, potentially mediated by mirror neuron processes (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005; Williams, Whiten, Suddendorf, & Perrett, 2001). According to Hatfield and colleagues (Hatfield et al., 1994; Hatfield, Cacioppo, & Rapson, 1993), the automatic and non-affective facial matching is assumed to be the foundation of emotional contagion, whereas others (Cacioppo, Martzke, Petty, & Tassinary, 1988; Dimberg, 1997; Winkielman & Cacioppo, 2001) discuss facial mimicry responses as resulting from emotional processes rather than from non-affective reflexive routes.

Facial mimicry, indeed, may apparently be influenced by the social and emotional context. Firstly, Dimberg (1997) has illustrated that facial responses may in fact occur as reactions to non-facial affectively valenced stimuli as well. Further, emotion induction is able to accordingly modify facial mimicry responses (Moody et al., 2007), emphasizing the relation of congruent facial responses to the action tendencies associated with the emotional reaction towards the stimulus and not primarily congruent with the emotion of the presented face. Social factors, like a competitive situation, may in fact evoke counter mimicry[4] (Lanzetta & Englis, 1989), even if the context is only non-consciously primed (Weyers, Mühlberger, Kund, Hess, & Pauli, 2009). Other research has demonstrated significant influences of previous liking of or attitudes towards the stimulus material on facial mimicry (McHugo, Lanzetta, & Bush, 1991; McIntosh, 2006). However, it is important to note that these findings often occur in longer time windows after stimulus

---

[4] Counter-mimicry describes the lack of mimicry responses or the expression of contrasting facial displays towards the respective emotional faces and is regarded to occur from a reduction or absence of affiliation through social, emotional, or motivational variables such as group membership, disliking the displayed other, or competition situations (e.g. see Bourgeois & Hess, 2008; Lanzetta & Englis, 1989).

onset (5-15 seconds to 2 minutes after stimulus presentation) than those used for other explorations of rapid facial mimicry responses (see Dimberg, Thunberg, & Elmehed, 2000; McIntosh et al., 2006; Moody et al., 2007; usually ranging from 0-1000 to 0-2000 ms after stimulus presentation). This implies potentially different processes of facial mimicry responses, for instance a hardwired non-affective motor matching in early time frames, which later shifts to an involvement of higher-order social and affective mechanisms. However, Moody et al. (2007) presented evidence of an emotion-congruent component of facial mimicry within the early time frame of 0-1000 ms after stimulus exposure that obscured the direct matching of facial expressions. Hence, there is a need for addressing those different response windows when subsequently exploring facial mimicry responses.

Taken together, once probably serving an evolutionary advantage regarding physical survival, facial mimicry processes nowadays could be regarded as kind of "social glue", as a behavior repertoire facilitating bonding, affiliation, communication, and social cohesion. These responses may be explained by two potentially interacting mechanisms – a direct, non-affective, and automatic matching response and also rapid affective reactions generated by appraisal processes that may be altered by the social and emotional context. In case of a mediation of motor mimetic facial reactions by proposed mirror neuron processes (Gallese, 2003), the emotional state or social context may modify the activation of this supposed neuronal base.

Several researchers underpin the notion of facial responses being a purposeful communication tool and view emotional reactions as communicative acts (Bavelas, Black, Chovil, Lemery, & Mullett, 1988; Fridlund, 1991; Fridlund et al., 1990). Prominent investigations by Fridlund (1991; Fridlund et al., 1990) on audience effects showed facial EMG activity to vary with perceived sociality of the viewing context and not primarily with felt emotions regarding the presented stimulus material. This, however, addresses the unanswered question of directionality – is the social context affecting the emotional state of a person and that way altering the facial reactions or is the social context determining facial mimicry directly?

As Moody et al. (2007) discuss, more research is needed in order to investigate what non-affective components are involved in facial mimicry responses as well as what parts of the basic mechanism are attributable to mimetic matching and which are affective reactions, but additionally, future research ought to address the functional significance of rapid facial mimicry responses. This poses a couple of sensible questions – how particularly sensitive is facial mimicry to human faces? And how exactly are factors like stimulus intensity and familiarity influencing facial mimicry responses towards human and non-human faces?

As facial mimicry is functionally regarded as a facilitator of social interaction and interpersonal bonding, there are no benefits of mimicking non-human, unfamiliar faces, like robotic or schematic stimuli depicting emotional expressions compared to corresponding human emotional stimuli. However, if there is a clear distinction between automatic motor mimetic and socio-emotional processes, especially in early time frames, a corresponding matching of the displayed emotions should occur independent from face type. Accordingly, a face-like configuration displaying specific emotional features should suffice to induce rapid facial mimicry responses. This links nicely to the debate on the explanatory theories behind social responses towards non-human entities and social robots in particular that were discussed in Chapter 1. If specific human-like attributes should suffice in automatically evoking comparable socio-emotional responses, the facial mimicry paradigm seems ideally suited to explore the profoundness behind these and also offers a unique way of investigating the functionality of facial mimicry responses and its potential influencing factors in general.

Additionally, a comparative investigation between human and non-human, but humanlike, stimuli in this domain offers insight regarding another, rather methodological issue. Already, artificial agents are even functioning as sole stimulus material in human emotion research (see Blascovich et al., 2002; Likowski, Mühlberger, Seibt, Pauli, & Weyers, 2008; Schilbach et al., 2006). However methodologically practical, the validity of such applications is not entirely clear, since up until now there remains a need for comparative investigations whether the social and emotional responses towards humans and non-human agents are indeed stemming from the same (neuro-)psychological pathways, whether these processes are exactly the same and indeed interchangeable. As discussed in Chapter 1, physiological and neuropsychological comparative studies between human and non-human, but humanlike, emotional stimuli are sparsely represented in literature, especially when looking at robots. Investigations from Dubal et al. (2011), Rosenthal-von der Pütten et al. (2013), and Chaminade et al. (2010) are three of the few and although they underscored comparable processes, brain responses were less pronounced concerning robotic stimuli or not completely equitable for human and robotic stimuli. However, similarities in the detection of emotion in both human and robotic faces (Dubal et al., 2011) lead to the supposition that the human brain seems indeed equally sensitive to expressive emotional displays without them necessarily being encoded exactly like human faces. Going one step further and regarding not only the detection of emotion but rather the motoric reaction towards emotion – will facial mimicry reactions towards emotional stimuli differ between human, robotic, and schematic faces?

Echoing Nass and colleagues (e.g. Nass & Moon, 2000), as the most basal humanlike cues should already trigger congruent socio-emotional responses, a simple schematic face should fully suffice in generating corresponding facial mimicry responses. As robotic faces contain even more humanlike characteristics, they should equally evoke congruent mimicry reactions, which would also underline Epley et al.'s (2007; Epley, Akalis, et al., 2008; Epley, Waytz, et al., 2008) account on anthropomorphization and humanlikeness. Yet, according to the uncanny valley concept (Mori, 1970; Burleigh et al., 2013; Moore, 2012), robotic faces may cause less and/or dissimilar mimicry responses compared to either human or schematic faces, as they are located in-between the two categories of "clearly human" versus "clearly non-human" along the hypothetical scale of humanlikeness.

Hence, the first of the following two investigations (Study 3 and Study 4) was conducted to compare facial mimicry responses towards human, robotic, and schematic emotional faces, measured via EMG activity over *m. zygomaticus major* and *m. corrugator supercilii* during presentation of happy and sad human, robotic, and schematic faces. As mentioned above, these studies may offer novel findings regarding the profoundness and neuropsychological mechanisms of social responses towards non-human entities. Moreover, these findings can give first insights into what extent facial mimicry responses are automatic, non-affective, motor mimetic reactions, and whether factors like familiarity, humanlikeness, or prototypicality potentially alter corresponding facial responses.

The second investigation was conducted with regard to potential influencing factors of facial mimicry responses, specifically focusing on stimulus intensity and familiarity. This, in turn, also allowed for an exploration of novelty and familiarity issues regarding social responses towards humanlike robots. Effects of stimulus intensity on facial mimicry responses towards different types of emotional faces have not yet been examined, however, there exists some evidence that familiarity might positively affect facial EMG reactions of *m. zygomaticus major* (Winkielman, Hall, Fazendeiro, & Catty, 2006). Yet, that study measured facial reactions towards dot patterns and not emotional faces, obscuring the transferability of these results, as emotional facial expressions pose functionally more significant stimuli.

To explore whether potential differences in the responses elicited by human and robotic emotional faces could be diminished by altered stimulus characteristics, in a follow-up study, the setup from Study 3 was extended to include stimuli depicting strongly pronounced smiling human and robotic faces in addition to the happy and sad emotional displays of human and robotic faces used in Study 3.

Further, it was investigated whether familiarization with the depicted robots had an influence on the responses towards the robotic faces. Familiarization should lead to response facilitation, which could diminish potential differences between mimicry responses towards human and robotic emotional stimuli. Two forms of familiarization were realized – firstly, a short video displaying the robot was presented prior to EMG measurement, secondly, regarding a different sample, previous interaction with a humanlike social robot in the past was used as quasi-independent variable.

This study design, in addition to examining the influence of stimulus intensity and familiarization, also offers the exploration of potential interactions between these factors. Moreover, different time frames of facial mimicry responses were regarded exploratively, considering potentially varying influences of the dependent variables on early and later underlying processes of facial mimicry reactions.

3.2 Study 3 – Facial Mimicry towards Robotic and Schematic Faces

3.2.1 Methods

3.2.1.1 Participants and study design

Thirty-three right-handed women with a mean age of 24.79 years (*SD* = 5.42 years) took part in the study with a 3 (*face type*: human, robotic, schematic) × 2 (*emotion*: happy, sad) repeated-measures design. Five trials per stimulus category were randomly presented for 5000 milliseconds each. Facial muscle activity over *m. zygomaticus major* and *m. corrugator supercilii* was assessed via electromyography during stimulus presentation and functioned as dependent measure for facial mimicry responses.

Participants were recruited on campus of Bielefeld University and received either monetary compensation or partial course credit for their participation. Testing only women in this first study eliminated potential confounding with gender effects. Additionally, earlier findings indicate women to be more emotionally expressive than men, resulting in more pronounced, but not quantitatively different facial mimicry (Dimberg & Lundquist, 1990). All participants were fluent in German, healthy, and currently not on any medication that might affect emotional, cognitive, or motoric functioning.

3.2.1.2 Stimulus material

Thirty pictures of emotional facial expressions were used, consisting of five pictures per stimulus category (see Fig. 7). Depicted were static happy and sad female human, robotic, and schematic faces. Pictures of the human emotional expressions were color photographs taken from the NimStim stimulus set (Tottenham et al., 2009) with five different female faces each displaying the two emotional expressions. For the robotic pictures, color photographs of the Flobi robot head (Hegel et al., 2010; Lütkebohle et al., 2010) in its female version were taken, showing either a happy or sad facial expression. Distinct combination of different hair and lip colors resulted in five individual robotic faces each depicting the two emotional expressions. A previous study confirmed that the emotional expressions could be correctly attributed with detection rates ranging from over 83% to 99% (Hegel et al., 2010).

The schematic faces were designed as highly prototypical “smiley”-like stimuli with only the mouth region varying according to the emotional expression – an upward lip pattern indicating happy affect and a downward lip pattern signalizing sad affect. To produce five distinct schematic faces, the schematic stimuli were underlaid with five different colors. All stimulus pictures were matched in luminance and height.

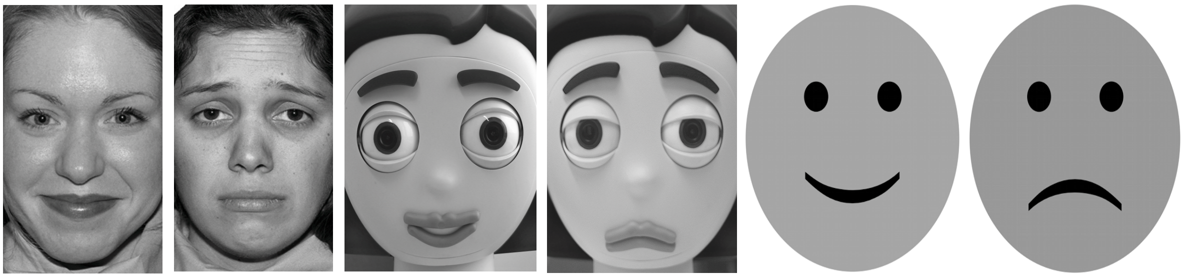

Figure 7. Exemplary stimulus pictures for the human, robotic, and schematic face types depicting happy and sad facial expressions. Copyright for robotic pictures: Frank Hegel, 2011.

3.2.1.3 Procedure

Participants were tested individually in a laboratory at Bielefeld University. Upon arrival, they were shown around the lab and informed about the experimental procedure. To obliterate the true purpose of the study in order to minimize specific facial reactivity effects and conscious facial movements, participants were told that skin conductance was measured instead of electromyographic activity. After signing a consent form, the EMG electrodes and an additional mock electrode on the left index finger (to allegedly measure skin conductance) were attached.

Participants were seated in a comfortable chair in an electrically shielded, dimly lit room. In order to reduce any distracting noise, they wore in-ear headphones playing pink noise throughout the EMG recording.

The stimulus pictures were randomly presented on a 19” CRT computer screen at eye level, approximately 75 cm away from the participants. Every trial consisted of a 3000 ms preparation period, in which they were asked to focus on a fixation cross in the middle of the screen. This was followed by 5000 ms of stimulus presentation, ensued by a variable inter trial interval of 5-15 seconds. Prior to the experimentation phase, participants completed five practice trials in order to get accustomed to the experimental procedure. After EMG recording, the electrodes were removed and participants were debriefed. No one reported having suspected the true purpose of the experiment.

3.2.1.4 Data recording, reduction, and analysis

Facial EMG was recorded bipolarly over the zygomaticus major and corrugator supercilii muscle sites (see Fridlund & Cacioppo, 1986) on the left side of the face (Dimberg & Petterson, 2000)[5] using standard Ag/AgCl electrodes (7 mm surface diameter). Previous to electrode attachment, the recording sites were cleaned with alcohol and prepared with abrasive paste. The surface electrodes were filled with Signa (Parker Laboratories Inc.) electrode paste.

For the recording, amplification, and filtering of the EMG signals, a QuickAmp amplifier and BrainVision Recorder software (BrainProducts GmbH) were used with a sampling frequency of 2000 Hz, a 20-500 Hz bandpass filter (24 dB/octave) as well as a 50 Hz notch filter (24 dB/octave). Subsequently, data were smoothed (8 Hz low-pass filter, 24 dB/octave), rectified, and baseline-corrected. The last 1000 ms prior to picture onset served as baseline.

Data were then segmented by respective stimulus category (combination of face type and emotional expression) and collapsed over trials. Only cases with > 50% of artifact-free trials per stimulus category were included in subsequent analyses.

For the two separate muscle sites, repeated-measures ANOVAs were calculated. Significant main effects were followed up using subsequent post hoc-comparisons and, in case of significant interactions, simple main effects analyses (Page, Braver, & MacKinnon, 2003). Uncorrected degrees of freedom are reported, however, Greenhouse-Geisser corrections were applied if necessary. If not indicated otherwise, two-sided significances are reported.

3.2.2 Results

3.2.2.1 Facial mimicry

In the following, the focus of interest lies on (a) mimicking of positive affect and (b) mimicking of negative affect. For the sake of clarity, the presentation of analyses is structured accordingly. First, analyses for mimicry reactions towards happy facial expressions are presented, quantified by EMG activity of both *m. zygomaticus major* and *m. corrugator supercilii*. Second, the corresponding analyses are presented for mimicry responses towards sad faces, again quantified by EMG activity of *m. zygomaticus major* as well as *m. corrugator supercilii*.

Mimicking of happy facial expressions

Regarding EMG activation of *m. zygomaticus major*, responses significantly differed along the presented emotional expression. Expectably, overall, happy faces led to more pronounced zygomaticus activity compared to sad faces, $F(1,32) = 10.74$, $p = .003$, $\eta_p^2 = .25$ (see Fig. 8). That is, perception of happy facial expressions led to an increase in activity of participants' own cheek muscle.

---

[5] Apparently, emotional facial reactions are more pronounced on the left side of the face (Dimberg & Petterson, 2000), which echoes the supposition of a greater involvement of the right-hemispheric motor cortex in the initiation of spontaneous emotional responses (Davidson & Hugdahl, 1995).

Concerning face type, a marginally significant main effect emerged, *F*(2,64) = 3.10, *p* = .062, η_p^2 = .09, which can be attributed to a significant difference between robotic and schematic faces, *F*(1,32) = 6.51, *p* = .016, η_p^2 = .17. This means that, independent from the emotional expression (happy vs. sad), robotic faces evoked overall weaker EMG responses in *m. zygomaticus major* than schematic faces.

Now, and of particular interest for the analyses of mimicry responses towards happy faces, this main effect of face type was qualified by a marginally significant face type × emotional expression interaction, *F*(2,64) = 3.04, *p* = .055, η_p^2 = .09. Follow-up analyses revealed that happy robotic faces elicited less pronounced zygomaticus activation than happy schematic faces, *F*(1,32) = 7.59, *p* = .010, η_p^2 = .19. Zygomaticus responses towards happy human faces lay in-between happy robotic and schematic faces, thus differing from neither one (all *p*s > .05).

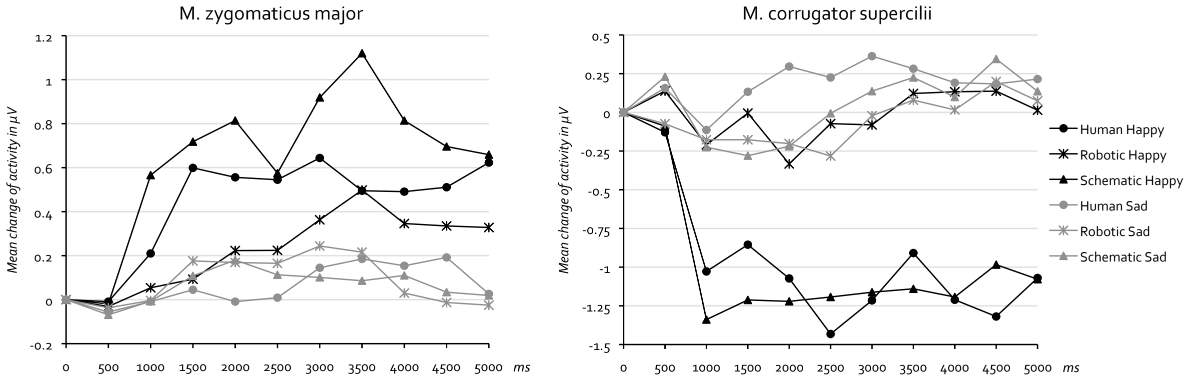

Figure 8. Mean electromyographic activity (in changes in µV) of *m. zygomaticus major* (left) and *m. corrugator supercilii* (right) in response to happy and sad human, robotic, and schematic faces.

Regarding the emotion-congruent activation, that is, a measurable response differentiation between happy and sad facial expressions of *m. zygomaticus major* for the different face types, only the human, *F*(1,32) = 7.00, *p* = .013, η_p^2 = .18, and schematic faces, *F*(1,32) = 8.50, *p* = .006, η_p^2 = .21, evoked significant emotion-congruent mimicry responses. That is, happy human as well as happy schematic faces elicited stronger zygomaticus activation than their sad counterparts. Yet, for robotic faces, mimicry responses did not clearly differ in congruence with the emotional expression, *F*(1,32) = 2.12, *p* = .155, η_p^2 = .06. Descriptively, an emotion-congruent differentiation of zygomaticus responses towards robotic faces was only detectable in the later time frames, starting around 3000 ms after stimulus onset (see Fig. 8).

With regard to EMG deactivation of *m. corrugator supercilii*[6], the three face types significantly differed from each other, as indicated by a main effect of face type, *F*(2,64) = 3.19, *p* = .048, η_p^2

---

[6] Note that *m. corrugator supercilii* is both susceptible for positive and negative affect and thus, presentation of happy facial expressions may lead to a deactivation of *m. corrugator supercilii* (see Larsen et al., 2003).

= .09. Closer inspection revealed that, overall, human faces differed significantly from robotic faces, $F(1,32) = 4.48$, $p = .042$, $\eta_p^2 = .12$, the latter provoking generally less corrugator responses, also in comparison with schematic faces, $F(1,32) = 6.75$, $p = .014$, $\eta_p^2 = .14$. Again, for the analyses of mimicry responses towards happy facial expressions, the face type × emotional expression interaction is of particular importance. Indeed, this interaction emerged as significant, $F(2,64) = 8.21$, $p = .001$, $\eta_p^2 = .20$. Simple effects analyses revealed that happy human faces led to greater corrugator deactivation compared to happy robotic faces, $F(1,32) = 9.37$, $p = .004$, $\eta_p^2 = .23$ (see Fig. 8). Also, happy schematic faces elicited stronger corrugator deactivation than happy robotic faces, $F(1,32) = 11.74$, $p = .002$, $\eta_p^2 = .27$. Thus, the perception of both happy human and happy schematic faces, but not of happy robotic faces, significantly decreased participants' brow muscle activity. No further significant effects were obtained regarding the responses towards happy facial expressions.

Mimicking of sad facial expressions

As mimicry responses towards sad faces are traditionally associated with an increase in activation of *m. corrugator supercilii*, results regarding *m. corrugator supercilii* are presented first: Concerning an activation of *m. corrugator supercilii,* predictably, sad facial expressions led to significantly more EMG activity than happy facial expressions, $F(1,32) = 10.57$, $p = .003$, $\eta_p^2 = .20$ (see Fig. 8). Regarding differential corrugator responses towards sad facial expressions depending on the specific face type, however, the three face types did not evoke significantly different responses, $F(1,32) = 0.48$, $p = .569$, $\eta_p^2 = .02$. However, emotion-congruent mimicry reactions for *m. corrugator supercilii* were only measurable for human, $F(1,32) = 12.50$, $p = .001$, $\eta_p^2 = .28$, and schematic faces, $F(1,32) = 9.96$, $p = .003$, $\eta_p^2 = .24$. Robotic faces on the other hand did not elicit significantly varying corrugator muscle responses for sad versus happy facial expressions, $F(1,32) = 0.09$, $p = .765$, $\eta_p^2 < .01$.

Regarding a deactivation of *m. zygomaticus major* in response to sad facial expressions, no differential effects concerning the face type were obtained, $F(1,32) = 0.07$, $p = .935$, $\eta_p^2 < .01$. No further significant findings emerged concerning mimicry responses towards sad facial expressions.

3.2.2.2 Additional analyses

Rating of humanlikeness and familiarity

An additional rating study assessed the subjective humanlikeness as well as the familiarity of the presented stimuli. Regarding humanlikeness, schematic faces were rated as significantly least humanlike, followed by robotic faces and human faces, which were regarded as most humanlike.

All pairwise comparisons differed significantly from each other (schematic vs. robotic: $t(20) = 3.33$, $p = .003$, $d_z = 0.73$; robotic vs. human: $t(20) = 12.31$, $p < .001$, $d_z = 2.67$; human vs. schematic: $t(20) = 15.81$, $p < .001$, $d_z = 4.53$), see Fig. 9.

A slightly different pattern emerged concerning the rated familiarity of the stimulus material. Expectably, human faces scored highest on this measure. Robotic faces were rated as least familiar and results for the schematic faces lay in-between the two other face types (human vs. schematic: $t(22) = 2.38$, $p = .027$, $d_z = 0.53$; schematic vs. robotic: $t(22) = 3.50$, $p = .002$, $d_z = 0.74$; robotic vs. human: $t(22) = 6.58$, $p < .001$, $d_z = 1.37$), see Fig. 9.

Happy-sad categorization task

To include an additional, conscious motoric reaction measure, response latencies during a happy-sad categorization task were assessed in a follow-up study ($N = 20$). Participants were required to quickly categorize the presented faces in a forced-choice reaction task as either depicting a happy or sad facial expression via a keypress. The pictures were presented in fully randomized order.

Of major interest was whether the three face types would differ regarding the categorization latencies. Results indeed revealed a significant processing advantage for schematic faces compared to human and robotic faces, which did not differ significantly from each other. Schematic faces could be categorized as significantly fastest (main effect *face type*: $F(2,36) = 6.01$, $p = .006$, $\eta_p^2 = .25$), see Fig. 9. Overall, the specific emotional expression did not significantly influence the response latencies, (main effect *emotion*: $F(1,18) = 0.02$, $p = .893$, $\eta_p^2 < .01$).

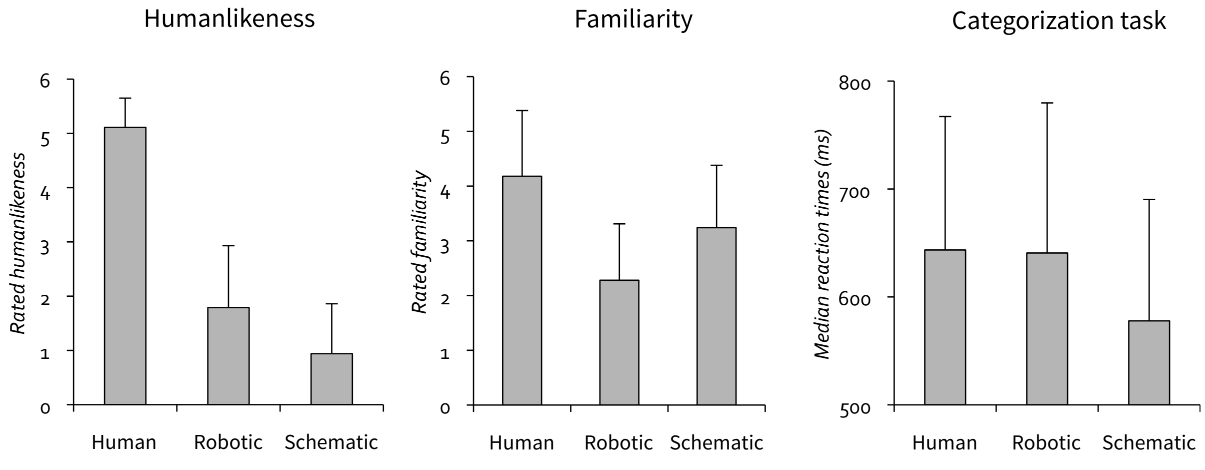

Figure 9. Descriptive results of the humanlikeness and familiarity rating as well as the categorization task. Error bars represent standard deviations.

3.2.3 Discussion

In order to explore whether facial mimicry responses would differentiate between human, robotic, and schematic faces, EMG activity over *m. zygomaticus major* and *m. corrugator supercilii* was measured during presentation of happy and sad human, robotic, and schematic emotional displays. Results revealed that facial mimicry reactions indeed differentiated between the different face types. Emotion-congruent facial muscle reactions in *m. zygomaticus major* and *m. corrugator supercilii* were present towards human and schematic faces. However, regarding robotic emotional faces, facial mimicry responses were significantly less pronounced. This was specifically apparent for happy facial expressions.

Further, happy and sad robotic faces did not cause differential *m. corrugator supercilii* activation as human and schematic faces distinctly elicited. Differentiating effects of stimulus valence on *m. zygomaticus major* responses for robotic faces were significantly smaller and occurred descriptively later than those towards human and schematic emotional expressions. These results allude to the aforementioned supposition that, in fact, the underlying processes of these low-level social responses towards non-human entities are not fully comparable with those towards humans. On one of the most basal response levels, findings regarding human (and schematic) faces appear not plainly transferable to robotic faces as well.

Concerning the additional subjective ratings, humanlikeness was most attributed towards human faces, followed by robotic faces, and then schematic faces. Regarding stimulus familiarity, human faces were rated most familiar, followed by schematic, and then robotic faces. An additional happy-sad categorization task revealed a significant processing advantage of schematic faces compared to human and robotic faces, the last two not differing from each other.

Most importantly, the presentation of robotic faces led to different facial mimicry responses than human or schematic faces. Regarding the facial mimicry reactions towards human faces as control responses, comparable reactions could only be detected for schematic faces. To explain these dissimilarities, both the diverging stimulus characteristics as well as the neuropsychological foundation of facial EMG will have to be taken into account.

Considering the lowest humanlikeness rating for schematic faces and also the shortest response latencies in the emotional categorization task (indicating that prototypicality rather than humanlikeness per se may be the crucial factor for these results), schematic faces likely profited from a significant processing advantage. Indeed, research has shown that facilitated stimulus processing even seems to be measurable via EMG. Easy-to-process stimuli appear to be able to elicit activation over *m. zygomaticus major* (Winkielman & Cacioppo, 2001). Additionally, Winkielman et al. (2006) propose prototypicality to be a fluency-enhancing variable and present a link between fluency, through prototypicality, and positive affect, that was measurable electromyographically. De Vries, Holland, Chenier, Starr, and Winkielman (2010) replicated this prototypicality-zygomaticus response link and present additional data depicting robust exposure effects of prototypicality on familiarity and liking. They argue that prototypicality might be the "strongest manipulation of familiarity or the underlying fluency" (p. 325). The results of the current study regarding the schematic and thus highly prototypical faces resonate well with this notion. Prototypical faces might likely elicit comparable facial mimicry responses to human faces through preferential processing.

On the other hand, robotic faces lay in-between human and schematic stimuli regarding prototypicality, and were rated least familiar. Considering facial mimicry, robotic faces led to overall less and later corresponding responses, which was specifically measurable for happy faces. Further, robotic faces did not elicit emotion-congruent muscle responses for *m. corrugator supercilii* (and only descriptively for *m. zygomaticus major*). They did, however, not differ from human faces on the categorization task regarding response latencies. Thus, while it may be safe to assume that the emotional display was indeed comparably recognizable, as the results from previous ratings and of the categorization task (see above) indicate, mimicry reactions towards robotic faces still appeared significantly dissimilar compared to human faces.

Stimulus unfamiliarity seems a potential explanatory factor as it might lead to opposing effects than fluency-enhancing variables like e.g. prototypicality. Whereas *m. zygomaticus major* activation for happy robotic faces was less pronounced and occurred later than for happy human or schematic faces, *m. corrugator supercilii* apparently did not show any differentiating effect regarding robotic faces, independent of stimulus valence. This is especially interesting as *m. corrugator supercilii* responses do not seem to be uniquely sensitive to negative affect alone. Indeed, it is both potentiated by pictures of negative affect and additionally inhibited by pictures of positive affect (Cacioppo et al., 1986; Lang, Greenwald, Bradley, & Hamm, 1993; Larsen et al., 2003)[7], as was measured for happy and sad human as well as schematic faces in the current study. Regarding the neurophysiology of *m. corrugator supercilii*, it is rather sparsely represented in the motor cortex and thus less prone to elaborate and precise voluntary motoric reactions, whereas *m. zygomaticus major*, as part of the well represented cheek and lower face region, is more involved in fine-tuned motor behavior as to sophisticated articulation, display rules and other voluntary motoric responses (Ekman & Friesen, 2003). Further, *m. corrugator supercilii* is mostly bilaterally innervated, in contrast to a greater contralateral innervation (Rinn, 1984) of *m. zygomaticus major*, which also promotes less fine motor control of *m. corrugator supercilli* compared to *m. zygomaticus major*. Hence, social and emotional influences on facial muscle activity might not simply result in inverted activation of

---

[7] The reverse does not hold for *m. zygomaticus major*, which rather shows quadratic effects of valence on muscle activity, and does not seem to be reciprocally activated by positive and negative affect (Larsen et al., 2003).

these two muscle sites, but rather in different patterns of activation. Activation of *m. zygomaticus major* is more likely to be influenced by social display rules than *m. corrugator supercilii*, whereas *m. corrugator supercilii's* receptiveness to both positive and negative affect may make it more susceptible for ambiguous or ambivalent stimulus material. Following its neurophysiological properties, Larsen et al. (2003) propose ambivalent stimuli to have antagonistic effects on the activity over *m. corrugator supercilii*, which should result in little change of activity compared to a baseline period. This quite accurately describes the pattern of EMG responses of *m. corrugator supercilii* towards robotic faces.

Regarding *m. zygomaticus major* responses, Larsen et al. (2003) argue positive affect to increase muscle activity whereas negative affect should inhibit, not necessarily decrease, muscle activity. Following this rationale, ambiguous stimuli may increase activity over *m. zygomaticus major*. Concerning *m. zygomaticus major* activity towards robotic faces, descriptively, an increase in muscle activity indeed occurs for both happy and sad robotic faces until 2500 ms after stimulus onset. Then, the muscle responses slightly differentiate along the valence dimension. The initial descriptive muscle activation for both robotic emotional expressions might depict the proposed activity response for ambivalent stimulus material on *m. zygomaticus major*. After 2500 ms stimulus duration, higher-order, potentially volitional motor responses of this "social" muscle might have come into play.

Taken together, facial mimicry responses towards robotic stimuli may be ascribed to potential stimulus inherent ambivalence of robotic faces. This could speak for a potential u-shaped function of prototypicality and humanlikeness on corresponding facial mimicry responses. Both high prototypicality and pronounced humanlikeness might be key factors in eliciting congruent mimicry responses. Accordingly, robotic faces may lie too much in-between the two categories of humanlikeness on the one side and prototypicality on the other side. If they were clearly prototypical, they would profit from response facilitation like schematic faces. If they were more humanlike so as physically hardly distinguishable from human faces, naturally, facial mimicry responses would likely not differ from those towards human stimuli. The results of the current study thus could portray physiological evidence of the uncanny valley concept (Mori, 1970), that is, of the revulsion response towards an artificial entity that, according to the uncanny valley hypothesis, occurs along the dimension of humanlikeness shortly before reaching full humanlikeness.

Interestingly, Dubal et al. (2011) showed that early brain processing of emotional robotic displays was not bounded to humanlike arrangements. Regarding the encoding of emotional expressions, robotic faces did not differ from human faces, despite cortically differing according to their "faceness" properties. This fits well to the present data of the categorization task, where the robotic emotional displays could be categorized equally fast as human emotional expressions. However, concerning an actual motoric reaction in response to a certain emotional display (in contrast to the mere recognition of said emotion), humanlikeness might play a decisive role after all. The results of the present study hence speak for no purely default process of facial mimicry, irrespective of the characteristics of the facial stimulus.

Nevertheless, there are clearly automatic and apparently non-conscious responses towards robotic faces, however, they are significantly less pronounced and occur later compared to human or schematic faces. Further, these reactions seem to be caused by stimulus ambivalence, possibly due to stimulus unfamiliarity as the subjective ratings of the stimulus material suggest. Also, specific physical stimulus characteristics might appear as potential influence factors, as the differentiating facial mimicry effects were specifically pronounced concerning happy robotic faces. Could a stronger smiling robot alter the present results and pull the facial mimicry responses towards robotic displays out of the "uncanny valley"? Green, MacDorman, Ho, and Vasudevan (2008) however argue that the tolerance for the range of acceptable facial proportions seems not strongly correlated with ratings of humanlikeness. Yet, they report that uncanny valley responses and feelings of creepiness are highest when participants are ambivalent about the presented humanness of the stimuli. Hence, in the subsequent study, the factors of stimulus familiarity as well as facial characteristics in terms of the intensity of the facial expression will be addressed as influencing mechanisms. More precisely, as the lower half of the face is an extremely informative characteristic regarding the facilitation of social interaction (see e.g. Adolphs, 1999; Haxby et al., 2000; Öhman, 2002), and the data of the current study show clear distinctions between the face types for the happy (and not for the sad) facial expressions, a manipulation of the intensity of the displayed smiles appears sensible, thus accentuating this social signal. This further allows for a more natural alignment of the mouth compared to eye size of the robotic stimuli, as the design of the Flobi robot incorporates disproportionally large eyes compared to its mouth size (see Hegel et al., 2010).

Regarding the debate on the nature and underlying processes of facial mimicry in general, the present study suggests that facial mimicry seems to comprise more than mere motor mimetic reactions. The present results rather allude to the position highlighting the importance of facial mimicry as a purposeful communication tool that might be particularly sensitive to human faces on the one hand or unambiguous, highly fluent face-like configurations on the other hand.

3.3 Study 4 – Influences of Expression Intensity and Familiarization on Facial Mimicry towards Robotic Faces

3.3.1 Methods

3.3.1.1 Research rationale

Potential influence factors of specific stimulus characteristics on the one side and stimulus familiarity on the other side were already discussed in relation to the apparent differences in facial mimicry towards human, robotic, and schematic faces found in the previous study. Thus, to explore whether the differences in the responses elicited by human and robotic emotional faces could be diminished by altered stimulus characteristics, namely expression intensity of the depicted smiles, the setup from Study 3 was adjusted to include stimuli depicting strongly pronounced smiling human and robotic faces in addition to the happy and sad emotional displays of human and robotic faces used in Study 3.

Further, the potential effects of familiarization towards the robotic stimuli were addressed. Familiarization should lead to response facilitation, which could diminish potential differences between mimicry responses towards human and robotic emotional stimuli. Hence, in the following, two forms of familiarization were realized – in a first experiment, a short video displaying the robot was presented prior to EMG measurement, whereas in a second experiment previous real-world interaction with a humanlike social robot was used as quasi-independent variable.

Apart from potential main effects of both stimulus intensity and familiarization, possible interactions between these factors were examined, as previous stimulus familiarization might alter the effects of stimulus intensity on facial mimicry responses towards robotic faces. Moreover, different time frames of facial mimicry responses were regarded exploratively, considering potentially varying influences of the dependent variables on early and later underlying processes of facial mimicry reactions.

3.3.1.2 Participants and study design

Forty-eight right-handed participants (40 women[8]) with a mean age of 25.1 years (SD = 3.93 years) took part in Experiment 1 with a 2 (within-factor *face type*: human, robotic) × 3 (within-factor *emotion*: happy, strongly happy, sad) × 2 (between-factor *familiarization*: naïve, familiarized) mixed-measures design. For Experiment 2, 98 right-handed participants (69 women[9]; mean age: 25.7 years, SD = 4.15 years) were recruited to take part in the same study design, except for the

---

[8] No significant effects including gender emerged (all *p*s > .05).

[9] No significant effects including gender emerged (all *p*s > .05).

difference of the between-factor familiarization being a quasi-independent variable of previous real-world interaction with a humanlike robot. Post hoc, participants were placed in the *previous experience* group when they reportedly had already experienced an interaction with a humanlike robot in the past. The *no previous experience* group consisted of participants who had not yet engaged with a humanlike robot, however, they were allowed to have previous knowledge of humanlike robots from TV or the Internet[10]. The determining factor for the difference between the two groups lay in the actual previous interaction with a humanlike robot.

As in the previous study, five trials per stimulus category were randomly presented for 5000 milliseconds each. EMG recordings of facial muscle activity over *m. zygomaticus major* and *m. corrugator supercilii* during stimulus presentation were assessed as dependent measures.

Participants were recruited on campus of Bielefeld University and received monetary compensation for their participation. All participants were fluent in German, healthy, and currently not on any medication that might affect emotional, cognitive, or motoric functioning.

3.3.1.3 Stimulus material

Five pictures per stimulus category were presented, depicting static happy, strongly happy, and sad female human as well as robotic faces (see Fig. 10). As before, pictures of the human emotional expressions were taken from the NimStim stimulus set (Tottenham et al., 2009) with five different female faces each displaying the three emotional expressions. The strongly happy emotional expressions resulted from more pronounced, mostly open-mouthed, smiling faces. For the robotic pictures, color photographs of the Flobi robot head (Hegel et al., 2010; Lütkebohle et al., 2010) were used, also depicting either a happy, strongly happy, or sad facial expression. Regarding the strongly smiling faces, the original pictures of the happy facial expression were digitally adapted using GIMP software in order to proportionally enlarge and also pull the edges of the mouth farther up and back.

A previous rating study ($N = 26$) confirmed that the pronounced smiling faces of both face types significantly differed according to their perceived emotional intensity from the respective normal smiles (normally versus strongly smiling robotic faces: $t(25) = 5.79$, $p < .001$, $d_z = 1.14$; normally versus strongly smiling human faces: $t(25) = 4.16$, $p < .001$, $d_z = 0.82$). All stimulus pictures were matched in luminance and height.

---

[10] This was assessed via a 4-item Guttman scale including „I know robots from the media.“, „I know humanlike robots from the media.“, „I have already interacted with a humanlike robot.“ and „I regularly work with humanlike robots on my job and/or in my free time.“ that was administered after the experimentation phase. None of the participants reported regular contact with humanlike robots either through their work or free time.

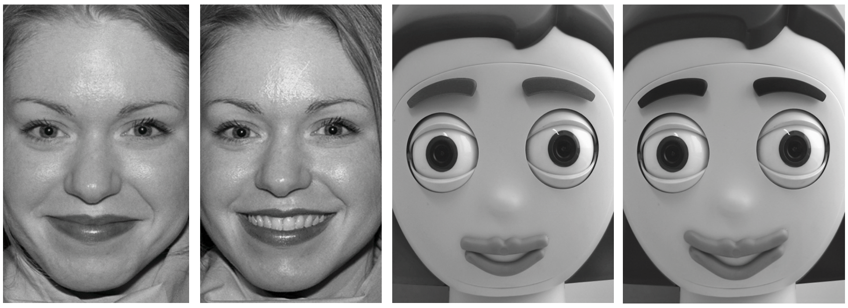

Figure 10. Exemplary stimuli of the normally (picture 1 and 3) and strongly (picture 2 and 4) smiling human as well as robotic face types.

3.3.1.4 Procedure

The experimental procedure and lab technique were similar to that of the previous study (see 3.2.1.3). Participants were tested individually in a laboratory at Bielefeld University. Again, they were told that skin conductance was recorded instead of muscle activity to minimize specific facial reactivity effects and conscious facial movements. In addition to the EMG electrodes in the face, a mock electrode was attached on the left index finger. To minimize distraction, in-ear headphones played pink noise throughout the EMG recording.

As experimental manipulation, in Experiment 1, prior to the start of the experiment, a short video clip (24 s duration) of the Flobi robot was presented to the video group, in which the robot was active and engaging in small-talk with a stranger[11]. By this perceptual familiarization, participants in this group (subsequently referred to as familiarized group) were thus given the chance to form a rough first impression of the actual robot that was used during the following experimentation phase. The other group (subsequently referred to as naïve group) did not watch a previous video.

Concerning the quasi-independent variable in Experiment 2, participants of both the previous experience and no previous experience group (subsequently referred to as familiarized and naïve group, respectively) did not receive any information regarding the robotic stimuli before the start of the experiment. The only difference between these two groups was the actual interaction with a humanlike robot in the past.

Prior to the experimentation phase, participants completed five practice trials in order to get accustomed to the experimental procedure. During the experimentation phase, stimulus pictures were presented in fully random order, where every trial consisted of a 3000 ms preparation period, in which participants focused on a fixation cross in the middle of the screen, followed by 5000 ms of stimulus presentation and a variable inter trial interval of 5 to 15 seconds.

---

[11] The conversation started with the person asking the robot about the current time, and a short exchange followed about where the person was headed next. The robot then encouraged the person to return later.

When participants were debriefed after the experiment was finished, none reported having suspected the true purpose of the study.

3.3.1.5 Data recording, reduction, and analysis

Following the procedures from the previous study, facial EMG was recorded bipolarly over the zygomaticus major and corrugator supercilii muscle sites (see Fridlund & Cacioppo, 1986) on the left side of the face (Dimberg & Petterson, 2000) using standard Ag/AgCl electrodes (7 mm surface diameter, filled with Signa electrode paste). The recording sites were cleaned with alcohol and prepared with abrasive paste prior to electrode attachment.

For the recording, amplification, and filtering of the EMG signals, a QuickAmp amplifier and BrainVision Recorder software (BrainProducts GmbH) were used with a sampling frequency of 2000 Hz, a 20-500 Hz bandpass filter (24 dB/octave) as well as a 50 Hz notch filter (24 dB/octave). Subsequently, data were smoothed (8 Hz low-pass filter, 24 dB/octave), rectified, and baseline-corrected. The last 1000 ms prior to picture onset served as baseline. Data were then segmented by respective stimulus category (combination of face type and emotional expression) and collapsed over trials. Only cases with > 50% of artifact-free trials per stimulus category were included in subsequent analyses. Finally, due to the between-subjects manipulation of this study, all data underwent within-subjects z-standardization.

For the two separate muscle sites, repeated-measures ANOVAs were calculated. Significant main effects were followed up using subsequent post hoc-comparisons and, in case of significant interactions, simple main effects analyses (Page et al., 2003). Uncorrected degrees of freedom are reported, however, Greenhouse-Geisser corrections were applied if necessary. If not indicated otherwise, two-sided significances are reported.

Regarding different time frames, results are presented for the global interval of the 0-5000 ms stimulus period. Additionally, analyses were calculated exploratively for the early time frame of 0-2000 ms after stimulus onset and the late time frame of 2000-5000 ms. For reasons of clarity and comprehensibility, the results of these analyses will only be appended in cases where they do not match the pattern of results of the global stimulus period, that is, when results in either the early or later time frame did not mirror the overall pattern of findings.

3.3.2 Results

The following sections comprise the findings of both experiments (Experiment 1: Video familiarization; Experiment 2: Real-world familiarization). Results regarding mimicry reactions towards happy facial expression are reported first, followed by the equivalent analyses for mimicry responses towards sad facial expressions. Similar to Study 3, mimicry reactions are quantified via EMG activity of both *m. zygomaticus major* and *m. corrugator supercilii.*

3.3.2.1 Experiment 1 – Video familiarization

Mimicking of happy facial expressions

Concerning *m. zygomaticus major* responses, the cheek muscle involved in pulling the corners of the mouth out- and upwards, expectably, happy facial expressions led to significantly stronger muscle activation than sad faces (significant main effect of emotion, $F(2,92) = 17.05, p < .001, \eta_p^2 = .27$). Globally, both normally and strongly smiling faces led to significant increases in zygomaticus activity compared to sad faces, $F(1,46) = 31.37, p < .001, \eta_p^2 = .41$, and did not differ from each other, $F(1,46) = 0.18, p = .671, \eta_p^2 < .01$ (see Fig. 12).

Furthermore, human faces led to overall more zygomaticus activity compared to robotic faces, as a significant main effect of face type emerged, $F(1,46) = 6.28, p = .016, \eta_p^2 = .12$. This means that, independent from emotional expression (happy vs. sad), robotic faces evoked overall weaker EMG responses in *m. zygomaticus major* than human faces. With particular attention to happy facial expressions, subsequent analyses revealed that globally, both normally as well as strongly smiling human faces led to significantly more zygomaticus activation compared to normally and strongly smiling robotic faces (effect of face type for normally smiling faces: $F(1,46) = 6.25, p = .016, \eta_p^2 = .12$; effect of face type for strongly smiling faces: $F(1,46) = 4.31, p = .043, \eta_p^2 = .09$). Interestingly, this variation between the human and robotic face type for normally smiling faces only occurred in the early time frame, $F(1,46) = 11.26, p = .002, \eta_p^2 = .20$, whereas the differentiation between human and robotic stimuli regarding the strongly smiling faces emerged as significant only in the later time frame, $F(1,46) = 4.64, p = .036, \eta_p^2 = .09$.

Intriguingly, the significant main effect of face type varied depending on previous familiarization. The significant difference between human and robotic faces regarding zygomaticus activity only occurred for the naïve group that did not watch a previous video displaying the robot, $F(1,46) = 5.86, p = .019, \eta_p^2 = .11$ (see Fig. 11). More specifically, in the naïve group, normally smiling human faces led to greater muscle activation than normally smiling robotic faces, $F(1,46) = 10.43, p = .002, \eta_p^2 = .18$. For the familiarized group, however, human faces overall did not significantly differ from robotic faces regarding the activation of *m. zygomaticus major*, $F(1,46)$

$=1.26$, $p = .267$, $\eta_p^2 = .03$. Yet, only for the late time frame of 2000-5000 ms and specifically regarding strongly smiling faces, human faces in fact evoked significantly larger zygomaticus responses than robotic faces in the familiarized group, $F(1,46) = 4.68$, $p = .036$, $\eta_p^2 = .09$.

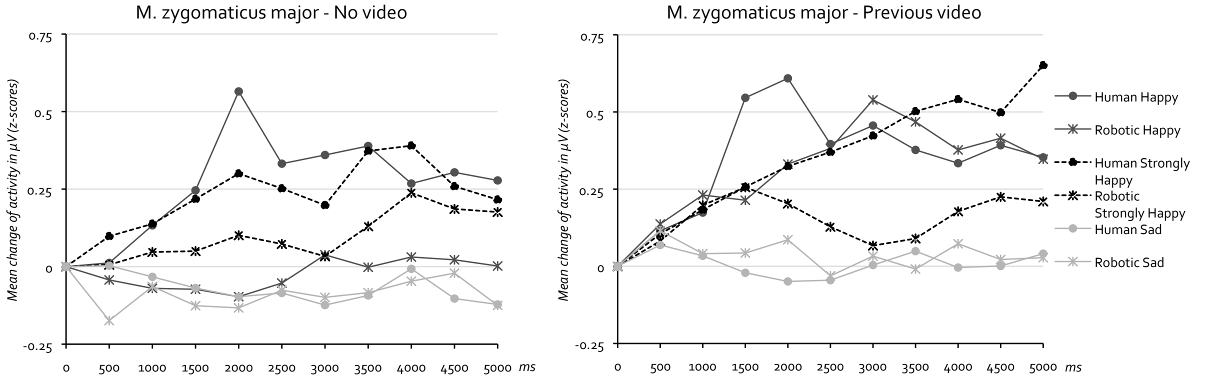

Figure 11. Mean activity of *m. zygomaticus major* (in changes in µV) of naïve (left) and familiarized (right) participants in response to human and robotic facial expressions.

Concerning the overall effects of stimulus familiarization on zygomaticus responses towards robotic faces, mimicry responses clearly differed in congruence with the emotional expression for the familiarized group, $F(2,92) = 5.72$, $p = .005$, $\eta_p^2 = .11$, however, they produced less discernable differences for the naïve participants, $F(2,92) = 2.49$, $p = .089$, $\eta_p^2 = .05$.

Interestingly, concerning potential effects of stimulus familiarization on the differentiation between normally and strongly smiling robotic faces, indeed, varying zygomaticus responses between the two familiarization groups were measured, $F(1,46) = 5.76$, $p = .021$, $\eta_p^2 = .11$. This effect can further be broken down into a marginally significant difference between normally and strongly smiling robotic faces for the familiarized group, $F(1,46) = 3.93$, $p = .054$, $\eta_p^2 = .08$, which did not occur for the naïve group, $F(1,46) = 1.99$, $p = .165$, $\eta_p^2 = .04$. That is, individuals who were previously familiarized with the robots showed differential mimicry reactions as a function of the intensity of the emotional expression, while participants without previous familiarization lacked this differentiation. However, the differentiation between normal and strongly smiling robotic faces in the familiarized group presents a pattern of increased activity for the normal smiling robotic faces compared to the strongly smiling robotic faces. Stimulus familiarization thus led to a differentiation between the two happy robotic facial expressions, however, it emerged in favor of the normally smiling robotic faces. Yet, this response pattern seems to be time-dependent, as it only emerged in the late time frame of 2000-5000 ms (interaction contrast of normally versus strongly smiling faces depending on familiarization: $F(1,46) = 5.36$, $p = .025$, $\eta_p^2 = .10$; normally versus strongly smiling faces familiarized group: $F(1,46) = 4.82$, $p = .033$, $\eta_p^2 = .10$; naïve group: $F(1,46) = 1.16$, $p = .287$, $\eta_p^2 = .03$) and was not observable for fast zygomaticus responses up until 2000 ms (interaction contrast of normally versus strongly smiling faces depending on familiarization: $F(1,46) = 2.75, p = .104, \eta_p^2 = .06$).

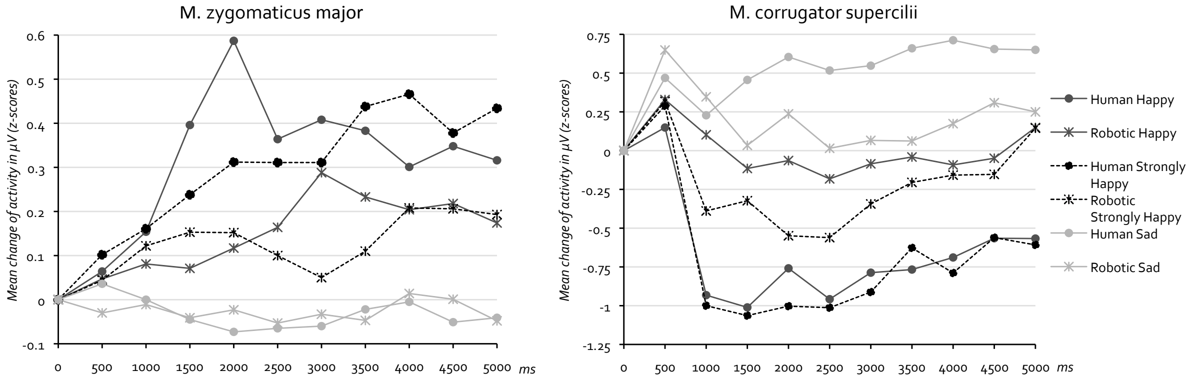

Figure 12. Overall EMG activity (as mean change of activity in µV) of *m. zygomaticus major* (left) and *m. corrugator supercilii* (right) in response to human and schematic faces presenting normally and strongly smiling as well as sad facial expressions.

Regarding *m. corrugator supercilii* responses, the inner brow muscle involved in frowning, which responds with activation towards negative affect and deactivation towards positive affect, globally, a significant deactivation during the presentation of happy faces was measured (significant main effect of emotional expression, $F(2,92) = 27.46, p < .001, \eta_p^2 = .37$). Both normally smiling and strongly smiling faces differed significantly from sad facial expressions, $F(1,46) = 51.80, p < .001, \eta_p^2 = .53$, however, overall, the two happy face types did not differ from each other, $F(1,46) = 1.11, p = .298, \eta_p^2 = .02$ (see Fig. 12).

Additionally, there was a main effect of face type, that is, overall corrugator responses towards human faces differed significantly from those towards robotic faces, $F(1,46) = 6.61, p = .013, \eta_p^2 = .13$. Yet, this effect was less pronounced in the later time frame, $F(1,46) = 3.64, p = .063, \eta_p^2 = .07$. As the face type × emotional expression interaction emerged as significant, $F(2,92) = 12.33, p < .001, \eta_p^2 = .21$, follow-up analyses revealed that both normally smiling human faces and strongly smiling human faces elicited significantly more corrugator deactivation than their robotic counterparts (effect of face type for normally smiling faces: $F(1,46) = 16.88, p < .001, \eta_p^2 = .27$; for strongly smiling faces: $F(1,46) = 7.75, p = .008, \eta_p^2 = .14$).

Interestingly, the previous main effect of face type varied as a function of prior familiarization, as corrugator responses towards human faces significantly differed from robotic faces only for naïve participants that did not watch a previous video of the robot, $F(1,46) = 6.58, p = .014, \eta_p^2 = .13$ (effect of face type for familiarized group: $F(1,46) = 1.15, p = .289, \eta_p^2 = .02$), see Fig. 13. More specifically, the previous significant difference between strongly smiling human and robotic faces was only present for naïve participants who did not watch the prior video displaying the robot, $F(1,46) = 7.77, p = .008, \eta_p^2 = 0.14$ (familiarized group: $F(1,46) = 1.32, p = .256, \eta_p^2 =$

.03). Additionally, the previous main effect of face type for naïve participants seems to be time-sensitive, as it was not measurable in the later time frame, $F(1,46) = 2.30$, $p = .136$, $\eta_p^2 = .05$. Regarding the effects of stimulus familiarity on the corrugator activity towards robotic faces, overall, the emotional expression only significantly differentiated responses accordingly when participants were familiar with the robot, $F(2,92) = 3.77$, $p = .027$, $\eta_p^2 = .08$, but not for the naïve participants, who did not watch the previous robot video, $F(2,92) = 0.80$, $p = .452$, $\eta_p^2 = .02$. Again, however, this congruent emotion differentiation in the familiarized group was largely due to early muscle responses, as the effect vanished in the later time frame of 2000-5000 ms after stimulus onset, $F(2,92) = 1.80$, $p = .172$, $\eta_p^2 = .04$.

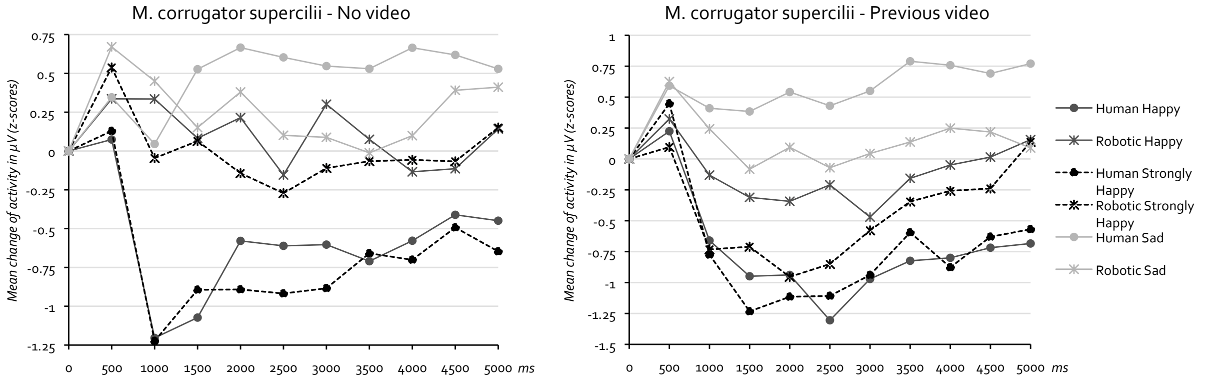

Figure 13. Mean activity of *m. corrugator supercilii* (in changes in µV) of naïve (left) and familiarized (right) participants in response to human and robotic facial expressions.

Indeed, again in the early time frame, normally smiling and strongly smiling robotic faces significantly differed from each other for participants in the familiarized group, $F(1,46) = 4.52$, $p = .039$, $\eta_p^2 = .09$, measurable as stronger deactivation for strongly smiling faces compared to normally smiling faces. Apparently, stronger corrugator deactivation for happy human compared to happy robotic faces may occur in case of unfamiliarity, but may be mitigated by more pronounced smiling of the robot when the robot is also familiar.

Mimicking of sad facial expressions

Regarding the mimicking of sad facial expressions, overall, *m. corrugator supercilii* activity was significantly higher in comparison to happy facial expressions, $F(2,92) = 27.46$, $p < .001$, $\eta_p^2 = .37$ (see above). Additionally, sad human faces led to significantly more pronounced corrugator muscle activation than sad robotic faces, $F(1,46) = 4.86$, $p = .036$, $\eta_p^2 = .09$ (see Fig. 12). Taking previous familiarization into account, this effect of face type seems to be attributable to the familiarized group, $F(1,46) = 3.96$, $p = .052$, $\eta_p^2 = .08$, whereas the differences in muscle activation towards the sad human versus sad robotic faces were not measurable for the naïve participants, $F(1,46) = 1.14$, $p = .291$, $\eta_p^2 = .02$ (see Fig. 13). However, the differentiation between sad human and robotic faces was only present in the later time frame of 2000-5000 ms after stimulus onset, $F(1,46) = 6.71, p = .013, \eta_p^2 = .13$, but not regarding the early muscle responses, $F(1,46) = 0.83, p = .366, \eta_p^2 = .02$.

Concerning *m. zygomaticus major* responses towards sad facial expressions, as mentioned above, overall, sad faces caused significantly less muscle activation and partly even muscle deactivation compared to happy facial expressions (significant main effect of emotion: $F(2,92) = 17.05$, $p < .001$, $\eta_p^2 = .27$; significant contrast of both normally and strongly smiling faces versus sad faces: $F(1,46) = 0.35, p = .558, \eta_p^2 = .01$), see Fig. 11 and Fig. 12. However, no face type specific muscle deactivation towards sad faces was measured, $F(1,46) = 0.01$, $p = .936$, $\eta_p^2 < .01$, likely caused by floor effects. That is, presumably no considerable muscle deactivation could occur following an already relaxed muscle during the pre-stimulus baseline period against which the muscle responses were corrected. No further effects were obtained regarding sad facial expressions.

Moderation by gender, empathy, or attitudes towards robots

Regarding the potential modulation of the found mimicry effects through variables like gender, personality traits like empathy (measured via SPF, Saarbrücker Persönlichkeitsfragebogen; Paulus, 2009), or personal attitudes towards robots (assessed via NARS, Negative Attitudes towards Robots Scale; Nomura, Suzuki, Kanda, & Kato, 2006), neither of the three emerged as significant covariate, all *p*s > .05.

3.3.2.2 Experiment 2 – Real-world familiarization

Mimicking of happy facial expressions

Regarding the activation of *m. zygomaticus major*, comparable to Experiment 1, happy facial expressions globally evoked significantly larger muscle responses compared to sad facial expressions (significant main effect of emotional expression, $F(2,192) = 23.80, p < .001, \eta_p^2 = .20$, and significant contrast of both normally and strongly smiling faces versus sad faces, $F(1,96) = 35.70$, $p < .001$, $\eta_p^2 = .27$), see Fig. 14 and Fig. 15. Concerning the two happy facial expressions, no measurable differences in zygomaticus activation were found towards normally smiling compared to strongly smiling faces, $F(1,96) = 0.55$, $p = .460$, $\eta_p^2 = .01$, mirroring the findings of Experiment 1.

Again, a significant main effect of face type emerged, that is, human faces led to overall more pronounced zygomaticus activation than robotic faces, $F(1,96) = 7.22, p = .008, \eta_p^2 = .07$. Further exploring a marginally significant face type × emotional expression interaction, $F(2,192) = 2.46, p = .088, \eta_p^2 = .03$, the differences between the two face types seem largely attributable to overall significantly stronger zygomaticus responses towards normally smiling human faces than towards normally smiling robotic faces, $F(1,96) = 10.68$, $p = .002$, $\eta_p^2 = .10$. Regarding the strongly smiling facial expressions, robotic and human faces tendentially differed depending on previous robot experiences, $F(1,96) = 3.03$, $p = .085$, $\eta_p^2 = .03$ (see Fig. 15). Participants who had not yet interacted with a social robot in the past showed significantly larger zygomaticus activity towards strongly smiling human faces compared to strongly smiling robotic faces, $F(1,96) = 9.78$, $p = .002$, $\eta_p^2 = .09$, however, the two face types did not cause differential muscle activation for strongly smiling faces when participants were already familiar with a real robot from previous experiences, $F(1,96) = 0.01$, $p = .936$, $\eta_p^2 < .01$. In Experiment 1, this pattern was not as clear. There, the differentiating effects of face type were attributed largely to greater muscle activation in response to normally smiling human compared to normally smiling robotic faces and this primarily regarding naïve participants.

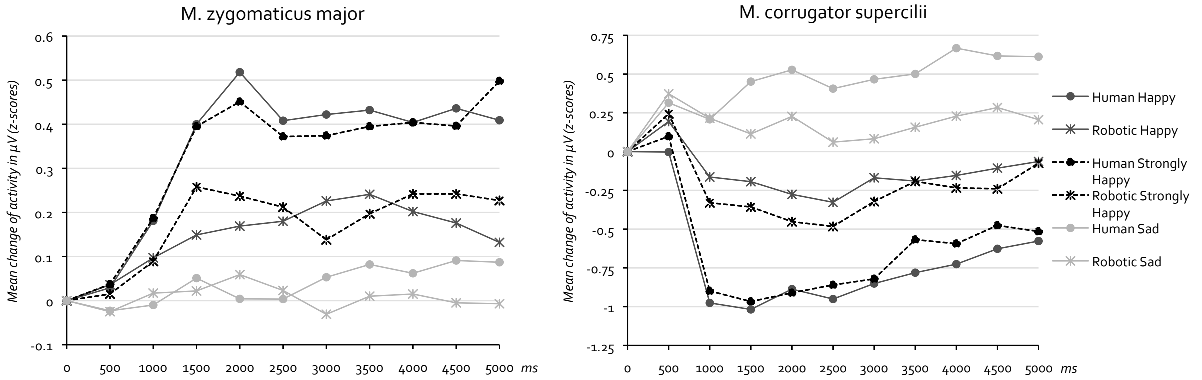

Figure 14. Overall EMG activity (as mean change of activity in µV) of *m. zygomaticus major* and *m. corrugator supercilii* in response to human and schematic faces presenting normally and strongly smiling as well as sad facial expressions.

Regarding the overall effects of previous robotic experiences on emotion-congruent zygomatic activity towards robotic faces, again comparable to the results of Experiment 1, the emotional expressions of the robotic stimuli led to globally congruent muscle responses only for the familiarized group, $F(2,192) = 7.87$, $p = .001$, $\eta_p^2 = .08$. However, naïve participants showed emotion-congruent zygomatic responses towards the robotic faces only in the later time frame of 2000-5000 ms, $F(2,192) = 3.14$, $p = .046$, $\eta_p^2 = .03$.

Furthermore, prior robotic experiences were indeed crucial for the potential differentiation between normally and strongly smiling robotic faces. Muscle activity did not significantly vary for naïve participants in response to normally versus strongly smiling robotic faces, $F(1,96) = 0.65$, $p = .424$, $\eta_p^2 = .01$, whereas it differed considerably for familiarized participants, $F(1,96) = 5.37$, $p = .023$, $\eta_p^2 = .05$, that is, stronger smiling robotic faces led to more muscle activity than normally smiling robotic faces in this group. However, the latter result was only present in the late time frame ranging from 2000-5000 ms after stimulus onset, $F(1,96) = 5.19, p = .025, \eta_p^2 = .05$. Note that this pattern of results runs inversely to that of Experiment 1, where the differentiation between normally and strongly smiling robotic faces emerged as stronger zygomaticus activation towards normally compared to strongly smiling robotic expressions.

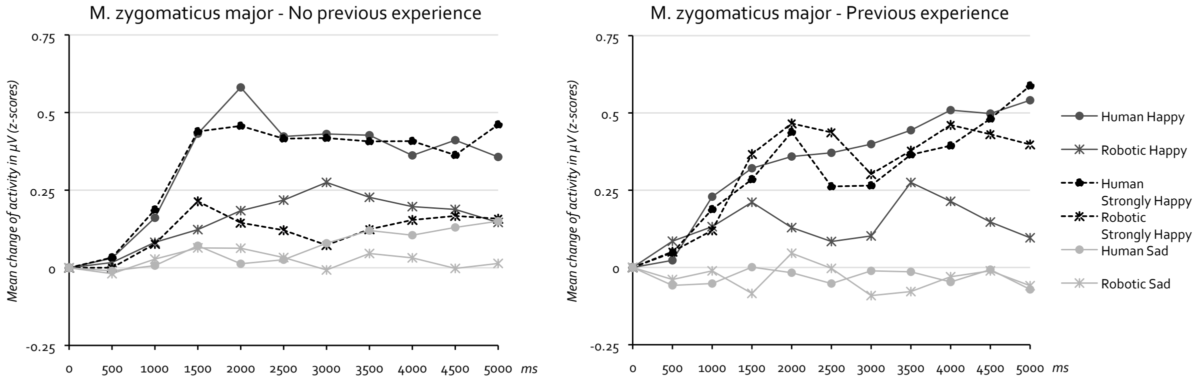

Figure 15. Mean activity of *m. zygomaticus major* (in changes in µV) of naïve (left) and familiarized (right) participants in response to human and robotic facial expressions.

Concerning *m. corrugator supercilii* activity, overall, happy faces led to significantly different responses compared to sad faces (significant main effect of emotional expression, $F(2,192) = 53.83, p < .001, \eta_p^2 = .36$, and significant contrast of both normally and strongly smiling faces versus sad faces, $F(1,96) = 99.18, p < .001, \eta_p^2 = .51$), that is, happy facial expressions led to a deactivation of *m. corrugator supercilii* (see Fig. 14 and Fig. 16). Globally, responses towards normally smiling faces did not differ from those towards strongly smiling faces, $F(1,96) = 0.23, p = .631, \eta_p^2 < .01$, echoing the findings from Experiment 1.

Follow-up analyses of a marginally significant main effect of face type, $F(1,96) = 3.13, p = .080, \eta_p^2 = .03$ (this face type main effect emerged as fully significant only in the early time frame, $F(1,96) = 7.77, p = .006, \eta_p^2 = .08$) as well as a significant face type × emotional expression interaction, $F(2,192) = 13.61, p < .001, \eta_p^2 = .12$, revealed that normally smiling human faces led to significantly stronger corrugator deactivation than normally smiling robotic faces, $F(1,96) = 16.37, p < .001, \eta_p^2 = .15$, again congruent with the results from Experiment 1.

Regarding the strongly smiling faces and breaking down the significant face type × familiarization interaction, $F(1,96) = 4.35, p = .040, \eta_p^2 = .04$, significant differences between the human and robotic face type emerged as dependent on prior familiarization, $F(1,96) = 7.48, p = .007, \eta_p^2 = .07$. Comparable to the pattern of results concerning familiarization effects in Experiment 1, naïve participants who had previously not experienced real-world interactions with a social robot showed significantly less corrugator deactivation towards the strongly smiling robotic faces than towards the strongly smiling human faces, $F(1,96) = 18.11, p < .001, \eta_p^2 = .16$, see Fig. 16. Conversely, no significant differences were obtained between strongly smiling human and robotic faces for participants that had previous robotic experiences, $F(1,96) = 0.30$, $p = .587$, $\eta_p^2 < .01$, mirroring the findings of the video familiarization in Experiment 1.

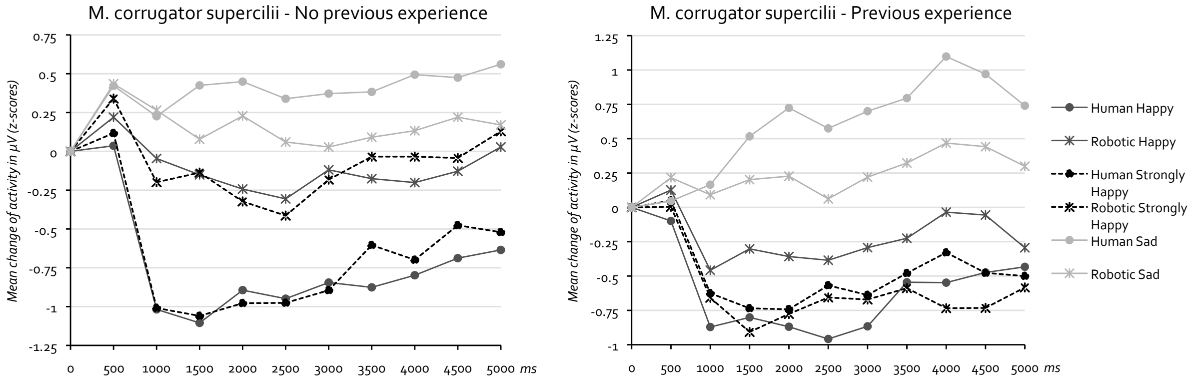

Figure 16. Mean activity of *m. corrugator supercilii* (in changes in µV) of naïve (left) and familiarized (right) participants in response to human and robotic facial expressions.

Despite these differences between the human and robotic face types as a function of prior familiarization, concerning the overall effects of previous robotic experience on corrugator activity towards robotic faces, the emotional expression of the robotic faces led to congruent corrugator responses in both the naïve, $F(2,192) = 3.15$, $p = .045$, $\eta_p^2 = .03$, as well as the familiarized group, $F(2,192) = 10.02$, $p < .001$, $\eta_p^2 = .10$. Yet, for naïve participants, this congruent emotional differentiation was only present in the early time frame, $F(2,192) = 4.12$, $p = .018$, $\eta_p^2 = .04$, and not significant regarding the later period of 2000-5000 ms after stimulus onset, $F(2,192) = 2.04$, $p = .133$, $\eta_p^2 = .02$. These results are largely comparable to the findings of Experiment 1, there, however, naïve participants did not show any differentiating emotion congruent mimicry responses towards robotic faces.

Additionally, apart from a globally stronger corrugator deactivation for the two happy facial robotic expressions versus the sad robotic faces (naïve participants: $F(1,96) = 5.80$, $p = .018$, $\eta_p^2 = .07$, familiarized participants: $F(1,96) = 14.73$, $p < .001$, $\eta_p^2 = .13$), a significant differentiation between normally and strongly smiling robotic faces was only present for familiarized participants, $F(1,96) = 4.48$, $p = .037$, $\eta_p^2 = .05$ (naïve participants: $F(1,96) = 0.03$, $p = .852$, $\eta_p^2 = .01$), which is in line with the results of Experiment 1. Furthermore, this effect occurred to be time-dependent, as it only emerged regarding the later time frame of 2000-5000 ms after stimulus onset, $F(1,96) = 4.49$, $p = .037$, $\eta_p^2 = .05$.

Mimicking of sad facial expressions

Regarding the mimicking of sad facial expressions, *m. corrugator supercilii* activation was significantly higher in response to sad facial expressions compared to happy facial expressions (see above; significant main effect of emotional expression, $F(2,192) = 53.83$, $p < .001$, $\eta_p^2 = .36$, and significant contrast of both normally and strongly smiling faces versus sad faces, $F(1,96) = 99.18$, $p < .001$, $\eta_p^2 = .51$), see Fig. 14 and Fig. 16. More precisely, breaking down the significant face type × emotional expression interaction, $F(2,192) = 13.61$, $p < .001$, $\eta_p^2 = .12$, sad human facial expressions evoked more corrugator activation than sad robotic faces, $F(1,96) = 6.31$, $p = .014$, $\eta_p^2 = .06$, comparable to the results of Experiment 1. This difference between the two face types, however, was only obtained in the later time frame of 2000-5000 ms after stimulus onset, $F(1,96) = 7.90$, $p = .006$, $\eta_p^2 = .08$, again largely echoing the findings of Experiment 1.

Regarding *m. zygomaticus major* responses towards sad facial expressions, globally, muscle activity was significantly lower for sad faces compared to happy facial expressions (see above; significant main effect of emotional expression, $F(2,192) = 23.80$, $p < .001$, $\eta_p^2 = .20$, as well as significant contrast of both normally and strongly smiling faces versus sad faces, $F(1,96) = 35.70$, $p < .001$, $\eta_p^2 = .27$), see Fig. 14 and Fig. 15. Concerning potential differences between human and robotic faces, no significant effects of face type were measured with regard to sad facial expressions, $F(1,96) = 0.16$, $p = .690$, $\eta_p^2 < .01$. This pattern of results of *m. zygomaticus major* responses towards sad facial expressions mirrors the findings of Experiment 1. No further significant effects were obtained regarding sad facial expressions.

Moderation by gender, empathy, or attitudes towards robots

As in Experiment 1, the factors gender, empathy (assessed via SPF; Paulus, 2009), and attitudes towards robots (assessed via NARS; Nomura et al., 2006) did not contribute significantly to the obtained mimicry effects, all *p*s > .05.

3.3.3 Discussion and conclusion

The current study was conducted to address the effects of robot familiarization as well as specific stimulus characteristics of the happy faces, that is, the intensity of the displayed smile, on facial mimicry responses. Therefore, in the present experiments, besides the happy and sad facial expressions of human and robotic faces from the previous study (see 3.2.1.2), additional strongly smiling robotic and human faces were presented. Furthermore, in Experiment 1 of the current study, stimulus familiarity was manipulated via a prior video displaying the robot in interaction, which half of the participants watched and the other half did not. In Experiment 2, a quasi-independent variable of familiarization was utilized – the actual experience of having interacted with a humanlike social robot in the past.

The current results revealed that facial mimicry towards human faces again followed the expected pattern, that is, emotion-congruent responses for both *m. corrugator supercilii* as well as *m. zygomaticus major* were found. During the presentation of happy faces, a deactivation of *m. corrugator supercilii* and an activation of *m. zygomaticus major* occurred, whereas sad faces provoked an increase in corrugator activity and a decrease of zygomaticus activation. Regarding the two happy human face types, no differentiation between normally and strongly smiling faces was measured. Facial mimicry towards human faces seems not to be linearly influenced by the intensity of the displayed smiles. Additionally, the muscle responses remained largely constant over the stimulus period of 0-5000 ms after stimulus onset, as no differing patterns of results were detected regarding the early (0-2000 ms) and later time frames (2000-5000 ms).

Resonating with the results of the first study, robotic faces, however, indeed globally differed from human faces. Again, facial mimicry responses towards robotic faces were significantly less pronounced than towards human faces and did not differentiate as clearly according to the displayed emotion. However, the two postulated influence factors of familiarization on the one side (that is, e.g., previous real-world experience with a social robot) and enhanced stimulus intensity (that is, a more pronounced expression of the smile) on the other appear to be accountable for these differences.

Remarkably, the differentiation from human faces was significant only for the naïve participants without prior experience with social robots in Experiment 1 and 2. For participants that were either familiar with the robot through the previous video or that had already interacted with a social robot in the past, overall, mimicry responses were largely comparable to those towards human faces. Also, emotion-congruent muscle responses were more pronounced for the familiarized compared to the naïve participants. Specifically regarding a potential differentiation between normally and strongly smiling robotic faces, prior familiarization indeed significantly influenced corresponding muscle responses. Mimicry reactions towards strongly smiling robotic faces differed significantly from those towards normally smiling faces, but only for the familiarized and not for the naïve participants. However, the type of familiarization apparently affected the direction of this difference. Whereas strongly smiling robotic faces led to significantly more corrugator deactivation than normally smiling robotic faces in both familiarized groups (hence, evoking similar responses as towards human faces), the response pattern of zygomaticus activity towards strongly smiling compared to normally smiling faces actually differed for these two familiarized groups. In Experiment 1, where participants watched a previous video displaying the concrete robot that was used for the emotional expressions, zygomaticus activity was surprisingly larger in response to the normally smiling robotic faces compared to the strongly smiling faces, whereas this pattern was reversed for the familiarization group in Experiment 2 that had prior interaction experience with a social robot. Here, previous experiences with a robot led to more pronounced zygomaticus activation towards strongly smiling robotic faces than towards normally smiling faces.

Familiarization through actual experience thus evoked mimicry responses towards strongly smiling robotic faces equivalent to those towards their human counterparts. Note that this previous real-world robot interaction did not necessarily have to be with the Flobi robot, but simply any social robot anytime in the participant's past. The fact that the two manipulations of familiarization produced a different mimicry response pattern regarding the two facial muscles, emphasizes the impact of real interaction on the formation of an adequate mental model of the robot, or rather robots in general. Even though the Flobi robot was presented in a previous video, where not only the mere physical stimulus properties were perceptible but also motion and vocal cues as well as the robot's interactional abilities, this apparently only affected the emotion-congruent corrugator responses towards robotic faces accordingly. However, the "social" zygomaticus muscle reacted accordingly only when the familiarization relied on real interaction with any social robot in the past. Hence, mere perceptual familiarization (here through the video presentation) appears to be insufficient in eliciting fully comparable mimicry responses towards robotic as well as human faces for both muscle sites. Instead, the formation of a basal mental model of robots through actual interaction emerged as the crucial factor. Taken together, unfamiliarity seems to foster the deviating mimicry responses towards robotic faces, yet these effects may be attenuated or even eliminated by previous real-world robotic experience when jointly working together with enhanced stimulus intensity, that is, more pronounced smiling of the robot.

Particularly, regarding mimicry responses over the time course of 0-5000 ms stimulus duration, effects involving muscle reactions of *m. corrugator supercilii* largely occurred in the early time frame of 0-2000 ms after stimulus onset, whereas those effects concerning *m. zygomaticus major* emerged mostly in the later time frame of 2000-5000 ms. This resonates well with the known neurophysiological differences between these two facial muscles (Larsen et al., 2003; Rinn, 1984). *M. corrugator supercilii* is sparsely represented in the motor cortex and thus less likely to be involved in fine-tuned motoric responses, the conscious expression of socially nuanced reactions, or masking of affective expressions. Hence, contrarily to *m. zygomaticus major* responses, early corrugator activity seems to express the spontaneous, automatic, and undistorted responses towards the emotional faces, mostly affected by stimulus valence. The cortically well-represented lower face region on the other side allows *m. zygomaticus major* a much greater involvement in voluntarily controlled responses, masking of affective reactions, and expression of social display rules. However, these muscle responses might therefore occur not as fast and reflectively as those of *m. corrugator* *supercilii*. They might instead be indicative of the influence of social contextual variables in these later time frames.

The results of the current study allude nicely to this. In both experiments, the manipulation of familiarization resulted in an alignment of *m. corrugator supercilii* responses towards strongly smiling robotic faces to the respective responses regarding happy human faces. The laboratory-induced previous stimulus exposure as well as previous real-world interaction had comparable effects on the mitigation of differences between the robotic and human face type when appropriate stimulus intensity was given (strongly smiling robotic faces). However, regarding *m. zygomaticus major* responses, the type of familiarization had great influence on the differentiation of facial mimicry responses towards normally and strongly smiling human and robotic faces. A laboratory familiarization through a previous video presentation of the robot interacting with another person did not suffice but rather counteracted the alignment of mimicry responses towards strongly smiling robotic with those towards happy human faces. Only previous real-world experience resulted in comparable zygomatic mimicry responses, and this only when robotic faces were strongly smiling. Thus, a previous social situation with a robot as well as enhanced stimulus characteristics are needed in order to elicit comparable mimicry responses towards happy robotic and happy human faces in *m. zygomaticus major*. As Larsen et al. (2003) point out, valence may have a greater influence on zygomaticus activity in social situations (versus for instance a laboratory environment). The results of the previous two experiments may support this thesis and additionally raise the issue that the prior social context may have outlasting effects on zygomatic activity even without concrete communicative intent of the present situation. To what temporal extent this might pertain, however, remains a question for future research.

Regarding the laboratory-induced familiarization, the lack of congruent zygomaticus activation for strongly smiling robotic faces even though participants were familiar with the robot through the previous video manipulation might also indicate the proposed threshold effect in order for stimulus valence to affect *m. zygomaticus major* responses (Larsen et al., 2003). Whereas *m. corrugator supercilii* activity varies in a rather linear function of both positive as well as negative affect, regarding potentiation of *m. zygomaticus major* responses, a certain intensity threshold supposedly has to be overcome. The results of the current study may indicate an even more nuanced position. As noted before, the video familiarization, in contrast to real-world experience, apparently did not suffice in generating a general mental model of robots, but can rather be regarded as perceptual familiarization with the stimulus. Thus, in the absence of a mental model, as the strongly smiling robots in this context pose no ecologically significant interaction stimulus, no social contextual factors existed that could have led to pass the response threshold for these faces, resulting in potentiated zygomaticus reactions varying with stimulus intensity. Rather, participants apparently merely responded accordingly to the stimulus that was perceptually matching, that is, they only mimicked the normally smiling robotic faces as these were congruent with the robot they previously watched in the video. Perceptual familiarization through the video thus only accordingly affected *m. corrugator supercilii*. This again can be integrated nicely with the neurophysiology of the two muscle sites, as the "social" zygomaticus responses indeed should be more affected by these social contextual variables compared to *m. corrugator supercilii*, which is regarded to be linearly affected by stimulus valence and thus, strongly smiling faces actually elicited congruent facial mimicry responses in the video familiarization group.

Previous research on the fluency-affect-link has demonstrated that mere exposure of previously unfamiliar faces indeed enhances zygomatic activity towards these stimuli (Harmon-Jones & Allen, 2001). The current results might expand these findings to the extent that fluency, through familiarization, may affect the responses of *m. zygomaticus major* and *m. corrugator supercilii* separately. Corrugator activity may be affected by fluency more directly, hence the congruent deactivation responses towards strongly smiling faces for both familiarization manipulations. Contrarily, in accordance with affective-perceptual fluency accounts (Reber et al., 1998; Winkielman et al., 2000; Winkielman & Cacioppo, 2001) and the neurophysiology of *m. zygomaticus major*, familiarization may increase fluency, however, social, situational, and contextual factors may determine the attribution of perceived fluency to affiliation or the specific response criterion in question, i.e. mimicking a displayed facial expression. Hence, when no functional mental model of robots is available, enhanced fluency, through perceptual familiarization, will only affect responses towards perceptually matching stimuli regarding the muscle site that is incredibly susceptible to socially significant stimuli.[12]

Certainly, a potential issue regarding the current experiments lies in the generalizability of the results to other robotic platforms. As this remains an important question for the expansion of the current results, especially regarding robotic platforms varying in the amount of anthropomorphic cues, it is quite remarkable that the most human-comparable mimicry responses towards robotic faces were obtained in the group of participants that had previously interacted with any

---

[12] Note however, that regarding corrugator activity in response to sad facial expressions, robotic stimuli still caused significantly less corresponding mimicry compared to human faces in the later time frames, largely detached from potential familiarization effects. Hence, sad robotic faces may still possess more potential to elicit ambivalence-related responses (in contrast to happy robotic faces), potentially due to the fact that empathizing with a sad robot bears even less ecological significance than affiliation with a happy robot. Additionally, even previous familiarization with robots in most cases will not include presentation of or interaction with a sad robotic counterpart, decreasing the likelihood of incorporating sad affect into the mental model of the robot. This further alludes to the aforementioned close link of fluency and specifically positive affect. Further research may reveal the exact mechanisms behind fluency differentially influencing responses towards stimuli incorporating positive compared to negative affect, yet the current results regarding sad robotic faces are in line with the global account on fluency specifically linked to positive affect (e.g. see Reber et al., 1998).

robot and not necessarily the particular Flobi robot platform. This may already indicate that the current results are potentially quite transferable to other robotic platforms as well.

When integrating the results of the current two experiments with the first mimicry study, the following picture emerges – despite the fact that the human brain apparently responds similarly towards human and robotic emotion (Dubal et al., 2011), the actual reactions towards these emotional displays in the form of facial mimicry responses are not fully comparable to one another. When scaled along humanlikeness and prototypicality, robotic faces seemingly lie too much in-between these two extremes, thus producing ambivalence in naïve participants, as they differed significantly from facial mimicry responses towards human and schematic faces. Ambivalence in turn may predict the pattern of mimicry results quite well (Larsen et al., 2003). Focusing on what causes this ambivalence, respectively, which factors could diminish the differences in mimicry responses, specific stimulus characteristics, i.e. the intensity of the displayed emotion, and stimulus familiarization emerge as two potent explanatory variables. The current study indeed produced evidence that both factors play their respective role and that, when they work jointly, are able to diminish the differing mimicry responses towards human versus robotic faces.

Regarding the application of the present results in human-robot interaction, one implication that can be drawn from the previous studies is that the integration of robotic emotional expressions may offer a way to indeed produce genuine low-level social affiliative responses in humans. This appears extremely beneficial for the constitution and maintenance of natural, positive interactions and relationships with social robots. Further, to the author's knowledge, the current studies present the first evidence that the electromyographic recording of muscle activity over distinct facial muscle sites provides an effective indirect and unobtrusive measure of these profound and basal social responses towards social robots, that is, using facial mimicry as a means to investigate social reactions towards social robots. Future research will have to explore the applicability of EMG measurement of facial mimicry responses in actual interaction settings, yet the current method could present a valuable way to obtain ground truth data.

However, the present findings also highlight potent predetermining factors of facial mimicry towards social robots that should be taken into account when applying both robotic emotional expressions into human-robot interaction as well as utilizing facial mimicry as dependent variable for the evaluation of human-robot interaction. For naïve participants, even though other overt social reactions might occur, the initial presentation of a robot will likely elicit low-level ambivalence-related responses. These however can be mitigated by familiarization with the robot, where the most potent form appears to be real interactional experience. The current findings underline that familiarization with a robot already works even on the most basal response levels. Yet, additionally, appropriate stimulus characteristics have to be met. Depending on the degrees of freedom of the particular robot face, specific importance should be assigned to the design of the lower part of the robot's face as the current findings indicate. This further emphasizes the functional importance and ecological validity of the detection of changes in the mouth region (Öhman, 2002), which is apparently transferable to robotic faces as well. Only when the displayed smiles of the robot were more pronounced did the corresponding mimicry responses in *m. zygomaticus major* and *m. corrugator supercilii* match the responses towards the respective human faces. And only in cases of familiarization did emotion-congruent and emotion-differentiating mimicry responses towards robotic faces emerge.

Regarding the functionality of facial mimicry, the findings of the presented studies suggest that mimicry responses are, in fact, not purely automatic processes, but rather that increased ecological significance may accordingly enable congruent mimicry responses. Stimulus ambivalence may, despite of the recognizability of the depicted emotion, impair congruent facial mimicry reactions. This ambivalence may be related to the factors of humanlikeness on the one side and prototypicality on the other side, as robots seem to fall in the uncanny valley between these two extremes. Prototypicality can be regarded as fluency-enhancing variable, which was supported by corresponding reaction time data. Indeed, when increasing fluency, through familiarization, the differentiating effects between human and robotic faces were mitigated. However, specific stimulus characteristics additionally had to be met concerning the lower portion of the face, that is, enhanced robotic smiles were presented, as this was crucial for the emergence of congruent facial mimicry towards happy robotic faces. Importantly, real-world familiarization elicited different zygomaticus response patterns than laboratory familiarization, even though both apparently equally affected corrugator activity congruently. This again interferes with a purely automatic, non-conscious approach on facial mimicry (Chartrand & Bargh, 1999; Hatfield et al., 1993; Niedenthal et al., 2005; Williams et al., 2001). Indeed, the current findings underline the neurophysiological differences of the two respective muscle sites (Larsen et al., 2003; Rinn, 1984) and extend previous research in supporting proposed assumptions according to influences of stimulus ambivalence (Larsen et al., 2003). Further, the current findings may present an integration of the neurophysiological basis of facial mimicry reactions regarding these two muscles and the "sociality" of their respective responses, which appear to be related to different aspects of stimulus processing, specific stimulus characteristics, and the influence of contextual and fluency-related variables.

4. General Discussion

Previous literature has shown that overt social responses found in human-human interaction appear transferable to human-robot interaction as well (Bartneck et al., 2005; Bartneck & Hu, 2008; Eyssel & Hegel, 2012; Eyssel & Kuchenbrandt, 2012; Kahn, Kanda, Ishiguro, Gill, et al., 2012; Klamer & Allouch, 2010; Kuchenbrandt et al., 2011; von der Pütten et al., 2011), as far as adequate reference cues, e.g. humanlike appearance cues, are available. Moreover, these mechanisms seem to work mindlessly and automatically (Nass et al., 1994; 1997; Nass & Moon, 2000). People for example feel pity for robots that get hurt, show politeness reactions towards them, are able to include robots in their in-group, cooperate with them, attribute personality traits, and apply the same stereotypes towards them as towards other humans. However, a closer look reveals that the underlying processes behind these responses are, in fact, less clear and potentially not truly equitable for human and robotic stimuli (Chaminade et al., 2010; Dubal et al., 2011; Rosenthal-von der Pütten et al., 2013). Hence, the current work was designed in order to examine subtler, low-level social and emotional responses towards robots both indirectly and unobtrusively. Hereby, the alleged profoundness of social responses towards non-human agents as well as their preconditions were studied in four experiments. The first two studies utilized an applied setting, the social facilitation paradigm. In a mere presence design, the social influence of a present robot was assessed via performance-based measures and manifested significantly compared to an alone control condition. Furthermore, the performance effects of human and robotic presence were comparable and additionally even sustained when the robot was obviously switched off.

In these studies, it was shown that the social influence of social robots indeed appears extremely profound, speaking in favor of mindless social responses elicited by basal social cues (Nass & Moon, 2000) and ruling out other explanatory mechanisms like overt socially learned responses, demand characteristics, superficial behavior adaptations, or reactions directed at the human behind the machine. As the neural level builds the core of all ensuing social responses, thus, to further investigate the boundary conditions of these allegedly profound and low-level social reactions, the subsequent two studies explored facial mimicry responses towards robotic faces via electromyographic assessment of muscle activity over corresponding facial muscle sites. Interestingly, facial mimicry responses towards robots significantly differed from both those towards human and schematic faces. This was specifically apparent for happy facial expressions. Apparently, additional conditions apart from a face-like configuration or humanlike appearance attributes have to be met in order to elicit comparable low-level socio-emotional responses.

Stimulus ambivalence was discussed as a probable explanatory factor for the obtained results regarding robotic stimuli. To alter stimulus ambivalence, specifically focusing on familiarization with the perceived faces and concrete face characteristics, namely the intensity of the displayed smile, the final study investigated whether the previously found differences between facial mimicry towards human and robotic faces could be mitigated. Indeed, real-world familiarization through actual interaction with a social robot in the past in accordance with more pronounced smiles led to comparable mimicry responses towards robotic faces as towards human faces.

Hence, despite comparable overt social responses towards robots and other artificial agents, an acute examination of the neuropsychological basis of these responses revealed that the mechanisms behind these reactions are, in fact, not fully interchangeable. Presenting a robot to naïve participants will likely induce ambivalence and subsequent related responses. However, when meeting specific preconditions, these differentiating factors can indeed be mitigated. Intriguingly, this is quantifiable indirectly and unobtrusively via extremely delicate measurements, underlining the profoundness and early stages of these responses and their influence factors.

Generally, the present work is in line with Nass and colleagues (Nass & Moon, 2000; Reeves & Nass, 1996) concerning the apparent mindlessness of social responses towards non-human agents. However, regarding their proposition that a minimal cue would suffice in activating low-level social response schemes towards non-human agents (Nass et al., 1994), the current work highlights a more nuanced approach. Whereas behaviorally, specific cues like a face-like configuration or humanlike appearance might be fundamental and sufficient for the elicitation of overt social behavior towards non-human entities, concerning low-level responses that are measured indirectly, the interplay of specific preconditions presents a pivotal factor. Humanlikeness alone was not sufficient for the emergence of congruent facial mimicry responses towards robotic faces. Prior familiarization paired with an optimized facial configuration appeared as the needed link in order to produce appropriate basal responses. Hence, the application of indirect and unobtrusive forms of measurement proved a sensible and fruitful approach in investigating the boundary conditions of social responses towards non-human agents.

Additionally, the current work offers the opportunity to review the evolving practice of using non-human agents as sole stimulus material in human emotion research (Blascovich et al., 2002; Likowski et al., 2008; Schilbach et al., 2006). Indeed, the similarities in the detection of emotion in human as well as robotic faces (Dubal et al., 2011) lead to the supposition that the human brain is equally sensitive to expressive emotional displays without the necessity of them being encoded exactly like human faces. The current work extended this notion to the involvement of actual motoric reactions towards emotionally expressive faces. The results indicate that the mechanisms behind the responses towards emotional displays of human versus robotic faces are, in fact, not truly interchangeable per se. The application of non-human agents in human emotion research thus contains potentially methodological issues regarding the equalization of the underlying processes and findings from such studies should be reviewed with great care.

Analogously, concerning the growing amount of robot-assisted therapy and the application of social robots in other healthcare scenarios, e.g. nursing homes (Dautenhahn, 2007; Kanda, Hirano, et al., 2004; Tapus et al., 2007), the transferability of therapeutic concepts should be gauged cautiously as the effects of ambivalence-related responses towards the robotic platforms may be especially detrimental in these domains where affiliation with the robot is of special importance. Inevitably, a great deal of attention has to be paid regarding the appropriate design of the robot as well as offering adequate amounts of familiarization, as the present data suggest. Regarding the application of social robots in workplace scenarios, the current work raised the issue that even the mere presence of a robot already significantly affects human performance. Furthermore, switching off the robot could not simply attenuate this influence. The potential benefits of robotic presence thus should always be weighed against the evident performance impairments this presence is able to induce, and potential strategies to effectively eliminate the robot's influence should be considered (e.g. using mobile robotic platforms that can either be sent away or that can even detect on their own when they should withdraw).

In both paradigms that were utilized for the current work, the social facilitation paradigm as well as the facial mimicry paradigm, stimulus ambivalence may have been a prime moderator of the obtained findings. As long as no prior mental model for a robot exists, it will apparently be encoded as ambivalent. Subsequently initiated mechanisms may include monitoring responses, heightened arousal, distraction, increased situational awareness, etc. Thus, even a switched-off robot may be able to hamper human performance. Further, it may already affect social responses on the most basal level – unconscious, fast, invisible motoric matching reactions. Apparently, familiarization may counter the effects of ambivalence related to the robot. Yet, as the mimicry studies revealed, laboratory familiarization might not fully suffice in enabling the generation of a mental model of robots, compared to real-world familiarization through actual interaction, as differences in the mimicry responses congruent with their "social susceptibility" depending on the type of familiarization indicated.

While the current study paradigms were suited to indirectly and unobtrusively investigate the boundary conditions of presumably automatic social reactions towards social robots, as discussed before, they are methodologically not fully convenient for studying how long exactly the social influence of the robotic agent and also potential ambivalence-related effects last. Yet, as the current findings indicate, this appears all the more relevant, especially with regard to the various potential application domains of social robots and other non-human agents. Anecdotal evidence (unpublished personal account of this author) of a three-week isolation study including, among others, daily interaction with the Flobi robot, which was permanently active throughout the day, suggests that the permanent robotic presence was, in fact, experienced as rather unpleasant throughout the whole course of the study and that participants were also not able to blind out the presence of the robot, whereas they were able to ignore the permanent video surveillance. Thus, while the mimicry studies indicate that stimulus ambivalence might be reduced by prior familiarization, the *social influence* of robotic presence may indeed persist over longer periods of time. Yet, future research may develop appropriate measurements and paradigms to innovatively explore the temporal domains of these effects.

Perceived humanlikeness was previously discussed as the prime candidate for the elicitation of social responses towards non-human agents (Epley, Waytz, et al., 2008). While prior research underlines the importance of humanlike cues for the generation of adequate responses (Epley, Akalis, et al., 2008; Nass et al., 1994; 1997; Nass & Moon, 2000; Nowak & Biocca, 2003), recent publications indicate that humanlikeness alone does not suffice (or, contrarily put, may not be needed) in evoking comparable or interchangeable responses towards non-human and human agents (Chaminade et al., 2010; Dubal et al., 2011; Rosenthal-von der Pütten et al., 2013). Dubal et al. (2011) even question the necessity of choosing a humanlike appearance when designing and developing social robots. The findings of the current work might conflate these positions in so far that humanlikeness appears as essential for the activation of social response schemes, yet, other factors, like stimulus familiarity, might superimpose on these and thereby influence the concrete manifestation of the specific reaction. The findings of Dubal et al. focus on the comparable encoding of emotion, whereas the current work extends this line of thought to include actual reactions towards displayed emotions. Concerning the actual performance of a response towards non-human agents, humanlikeness (combined with familiarity) might in fact be crucial in entailing the needed ecological significance of said social response. However, as the mimicry studies showed, even though the robotic stimuli included humanlike cues, they still appeared to lie too much in-between the categories of "human" on the one side and "prototypical/schematic" on the other side to readily evoke the same responses as human or schematic faces. Yet, when tuning these humanlike cues (e.g. by pronouncing the smiles of the robot faces), together with additional familiarization, comparable congruent responses could be obtained.

As these results thereby might be regarded as psychophysiological evidence of the assumed uncanny valley effect (Mori, 1970), additionally, the methodology of measuring facial mimicry responses could be utilized in experimentally investigating the optimal balance of humanlikeness in social robots in order to obtain desired responses. Nevertheless, the current findings underline that, while readily comparable on a macrolevel, that is, regarding the behavioral measures obtained in the social facilitation experiments, several precautions have to be taken in order to achieve a comparable social response towards a robot than towards a human when looking at the microlevel. This poses the legitimate question whether it is sensible to copy appearance, behavior, interactional patterns, voice, etc. from humans when developing a social robot, as apparently, several thresholds have to be passed in order to elicit comparable results. It might as well be sensible to concentrate robotic design on the pure functionality of the robot regarding its application domain, as this would likely enable a fast formation of a suitable mental model that should be largely based on the content of its application and not confounded by nuances of its humanlike appearance. In fact, recent research supports this notion and indicates that the preference of a humanlike or a mechanical robot appears dependent on the concrete application domain (Goetz, Kiesler, & Powers, 2003; Hinds, Roberts, & Jones, 2004). Hence, a social robot in a workplace scenario might profit from user acceptance and potentially induce less performance effects when its humanlikeness and social cues are drastically reduced, whereas social robots in social application domains could benefit from the opposite. Yet, as far as the application of social robots is still largely unaccustomed, these hypotheses remain for ensuing research. Until then, humanlike cues, specific stimulus characteristics (e.g. sufficient stimulus intensity), and familiarization appear as relevant preconditions for the most basal social responses towards social robots.

Apart from the implications drawn from the current findings regarding the application of social robots, the present studies may additionally provide valuable inferences concerning the psychological basis of the used paradigms. First, the social facilitation effects appear particularly profound, as this work was the first to show that they are equally transferable to robotic presence and even occur when the robot is switched off. This underlines the arousal and cognitive process approaches as explanatory mechanism (especially regarding the social inhibition part of the paradigm), because the evaluative components that are discussed as third explanatory factor were largely eliminated in the second study. In accordance with the original proposition of Zajonc (1980), unpredictability may be a potent underlying factor, influencing both physiological arousal and attentional diversion. Unpredictability may be especially pronounced for the presence of a robot. This substantially links to the influence factors of the facial mimicry responses towards robots, as stimulus ambivalence, through unfamiliarity, may be regarded as potent modulator of the found effects.

Concerning the functionality of facial mimicry, the present findings indicate that mimicry responses are, in fact, not purely automatic motor mimetic processes. As facial mimicry is largely regarded as facilitator of social interaction, affiliation, and bonding, mimicking non-human, unfamiliar faces poses no ecological benefits compared to mimicking human faces. Indeed, robotic faces apparently induced significantly less corresponding facial mimicry in naïve participants. This argues against fully automatic mechanisms that would foresee a face-like configuration displaying specific emotional features as sole requirement for the emergence of congruent mimicry responses (Chartrand & Bargh, 1999; Hatfield et al., 1993; Niedenthal et al., 2005; Williams et al., 2001). Apparently, neither a mere perception-behavior link (Chartrand & Bargh, 1999)[13] nor the hypothesis of biologically given affect programs by which humans are predisposed to automatically react emotionally to expressive facial stimuli (e.g. Dimberg, 1997; Dimberg et al., 2000) can fully account for the current findings. Yet, the factor of fluency appears to play a decisive role, as facial mimicry responses in fact were equally pronounced for highly prototypical schematic faces as they were for human faces. Hence, stimulus fluency may conflate the alleged contrary positions as the current findings uncovered. Both familiarization as well as prototypicality present fluency-enhancing variables (de Vries et al., 2010; Winkielman et al., 2006). Highly schematic emotional displays (Study 3) as well as familiarization paired with optimized stimulus characteristics (Study 4) led to comparable mimicry responses towards robot faces as towards human faces. Thus, fluency may present one of the boundary conditions of facial mimicry responses. Highly fluent stimuli, like human or schematic faces may readily produce the supposed "automatic" responses, thus refining Dimberg's account (1997; Dimberg et al., 2000). This also echoes the position that human attention evolved to develop category-specific selection criteria (see Dubal et al., 2011; New, Cosmides, & Tooby, 2007), which emotional patterns may be a part of and subsequently evolved as a signal of their own. In case of reduced fluency (e.g. through unfamiliarity or ambiguous category membership), however, intermediary processes between perception and behavior may impair the facial mimicry responses (as additional attentional processing of the human versus non-human defining perceptual features may be required and also the range of influence of social or contextual factors increases). This allows for the occurrence of belated and overall reduced mimicry reactions and potentially even counter-mimicry.

Regarding the neurophysiology of the two respective muscle sites, *m. corrugator supercilii* and *m. zygomaticus major*, the current findings contribute affirmative evidence to Larsen et al.'s (2003) hypotheses concerning the impact of stimulus ambivalence on facial EMG. Moreover, the present data demonstrate differentiating susceptibility as to the sociality of the respective muscle site, which echoes the neurophysiological basis of *m. zygomaticus major* and *m. corrugator supercilii* (Rinn, 1984). Whereas corrugator activity already congruently varied towards robotic faces under any form of familiarization, as long as the robotic smile was more pronounced, zygomaticus activation differed for laboratory and real-world familiarization, alluding to the social significance of expressing congruent mimicry that indeed should be more accentuated in this "social" muscle. A perceptual familiarization only evoked zygomatic mimicry responses towards exactly matching faces (hence, towards the normally smiling faces), whereas the formation of a general mental

---

[13] However, in their later work, Chartrand, Maddux, and Lakin (2005) soften their account and acknowledge the existence of potential motivational moderators on facial mimicry responses.

model of robots through previous real-world interaction enabled ecologically meaningful mimicry responses, that is, an intensity-congruent zygomaticus differentiation.

Taken together, the presentation of robotic stimuli allowed for a muscle site-related distinction between fluency and the ecological significance of facial mimicry responses that would normally coincide when investigating mimicry responses solely towards human faces. As Ghazanfar and Turesson (2008) argued, robots present a potential scaffold for neuroscientific thought. An integration of additional affective neuropsychological concepts into the investigation of social responses towards social robots thus proved a fruitful approach both for the field of social robotics as well as related domains of social and biological psychology.

Finally, the application of indirect and unobtrusive non-interactional measures, like performance data for the social facilitation studies or physiological data in the mimicry experiments, was designed as sensible methodology to assess the alleged profoundness of social responses towards non-human entities and indeed emerged as successful. This is especially relevant as potential confounds through e.g. direct interaction, demand characteristics of the specific situation, and self-report biases, which are usually present in studies exploring overt social responses towards non-human agents, could be ruled out. Further, as all social reactions ultimately build on a neural architecture, potential similarities or differences on these lowest of response levels can be regarded as highly relevant to affect and shape subsequent general social responses, affiliation, attachment, and acceptance of non-human agents. The present work provides valuable insights as to the robustness of social responses towards social robots, the potential implications concerning their application as well as the underlying psychological mechanisms. Possibly coupled with the promising methodological approaches employed in the current work, future studies might embrace and extend the presented research.

The proposed mechanisms behind anthropomorphization appear suitable to merge with and extend the results presented in the current work. According to Epley and colleagues (e.g. Epley, Waytz, et al., 2008), anthropomorphization embodies a form of inductive reasoning based on representations of the self or humans in general. This, however, could counteract with the direct, fast, low-level responses obtained in the facial mimicry studies as anthropomorphization apparently requires some higher-order processing, like the application of activated knowledge and integration of new information with initially activated representations, the use of heuristics, etc. (Epley, Waytz, et al., 2008). Still, as the current findings show, mimicry responses were not completely unmediated by the presented face type. Factors like humanlike appearance cues or robot familiarization indeed affected these supposedly default processes. Therefore, the use of the presented methodology could be a vehicle to disentangle the conflation of effects of humanlikeness with those regarding anthropomorphization responses, as the further naturally is a primary component of the latter but the two may still initiate and comprise distinct processes. As Krämer (2008) already noted, the ascription of humanness or humanlike representations on the one hand and a mindless social response on the other hand might even be two distinct and independent processes.

Hence, an integration of the alleged determinants of anthropomorphization appears rewarding. Dispositional factors like the tendency to anthropomorphize (Chin, Sims, DaPra, & Ballion, 2006; Epley, Waytz, et al., 2008), empathy-related personality characteristics, specific attitudes towards robots or factors like social phobia may represent likely influence variables. Note however, that subjective empathy and attitudes towards robots did not significantly account for any variance in the current findings. Yet, an experimental manipulation of these variables might obtain diagnostically more conclusive evidence.

Additionally, the exploration of motivational factors appears sensible because of their ecological significance. The fundamental need for social connection, termed sociality motivation, may play a large part in the initiation of social responses towards non-human agents and research on anthropomorphization already showed that loneliness indeed affected anthropomorphization responses (Epley, Waytz, et al., 2008). Will a manipulation of the need for affiliation also lead to measurable effects on low-level responses, like facial mimicry towards human and robotic faces? The apparent susceptibility of facial mimicry responses towards factors like familiarization and specific stimulus characteristics may imply this notion. Ensuing, considering experimentation with a live robot, another intriguing question lies in the exploration of effects of the robot actually mimicking (or purposely not mimicking) the human's facial expression. To what extent will the robot's mimicking behavior in turn affect the need for affiliation and subsequent mimicry and anthropomorphization responses?

Research further highlighted the influence of effectance motivation (Epley, Waytz, et al., 2008; Eyssel et al., 2011) on anthropomorphization, that is, when the sense of understanding and control (either through dispositional or situational factors) of one's environment is weakened, people tend to anthropomorphize more. This links nicely to the influence of unpredictability and ambivalence, which was previously discussed in relation to both the social facilitation as well as the facial mimicry paradigm. An experimental manipulation of these variables in both presented paradigms appears sensible in order to make out the profoundness of these influence factors (via the facial mimicry paradigm, e.g. by manipulating the degree of humanlikeness or the amount of humanlike cues of the robotic faces, or else by using dynamic facial expressions or a real robotic counterpart with either congruent, random, or incongruent facial expressions) and its implications regarding the actual application of social robots (via the social facilitation paradigm, e.g. by manipulating the randomness and/or humanlikeness of its behavior in the presence conditions).

Lastly, an investigation of the temporal domain of these factors appears highly relevant. Manipulations of the length of robotic presence or repeated robotic presence may uncover the influence of potential habituation or sensitization over time. This certainly is not only of relevance for the possible duration of the impact of potential influence factors on social responses towards non-human entities, but also regarding the longevity of the mere social reactions themselves. The facial mimicry paradigm might even offer an opportunity of assessing different habituation or sensitization steps when looking at the effects of presentation position during a fixed repeated-exposure experiment. This might enable a concrete visualization of the chronology of potential anthropomorphization and other higher-order processes and how to differentiate these from unmediated basal social responses.

Further investigation of the aforementioned factors thus may help in sounding out the boundary conditions of this prominent concept, which equally touches domains in social robotics as well as social cognition. Combined with the findings and presented methodology of the current work, future research may be able to strike new paths in its exploration.

4.1 General Conclusion

The core findings of the current work regarding the profoundness and preconditions of social responses towards social robots can be subsumed as follows:

Social responses towards social robots can be measured indirectly and appear indeed as extremely profound. They neither require an explicitly social situation or context to emerge, nor do they depend on actual interaction with the robot. Furthermore, the social robot apparently does not even need to be active to evoke social responses. The mere presence of a social robot is able to induce congruent social responses and its social influence pertains even when it is switched off.

Regarding the question of how far this influence may reach, apparently, it even affects the most basal of response levels, here, the fast, unconscious, motoric matching of observed facial expressions. Yet, concerning the profoundness of these low-level reactions, the responses towards robots are not fully comparable to respective reactions towards humans. Thus, the underlying processes behind these, in fact, appear not completely equitable or transferable. Comparable low-level social and emotional responses will only emerge when several preconditions are met. These necessary preconditions are (a) a humanlike appearance of the robot, paired with (b) sufficient intensity, that is, unambiguity of facial expressions presented by the robot and, most importantly, (c) familiarity, that is, previous interaction experience with social robots in general.

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, *3*, 469–479.

Adolphs, R. (2004). Emotional vision. *Nature Neuroscience*, *7*, 1167–1168.

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, *5*, 163–180.

Aiello, J. R., & Svec, C. M. (1993). Computer monitoring of work performance: Extending the social facilitation framework to electronic presence. *Journal of Applied Social Psychology*, *23*, 537–548.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Bartneck, C. (2003). Interacting with an embodied emotional character. *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*, 55–60.

Bartneck, C., & Hu, J. (2008). Exploring the abuse of robots. *Interaction Studies*, *9*, 415–433.

Bartneck, C., Chioke, R., Menges, R., & Deckers, I. (2005). Robot abuse – a limitation of the media equation. *Proceedings of the Interact 2005 Workshop on Abuse.*

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3–26.

Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R., & Mullett, J. (1988). Form and function in motor mimicry: Topographic evidence that the primary function is communicative. *Human Communication Research*, *14*, 275–299.

Berger, S. M., Hampton, K. L., Carli, L. L., Grandmaison, P. S., Sadow, J. S., Donath, C. H., & Herschlag, L. R. (1981). Audience-induced inhibition of overt practice during learning. *Journal of Personality and Social Psychology*, *40*, 479–491.

Bethel, C. L. (2009). *Robots Without Faces: Non-verbal Social Human-Robot Interaction* (Doctoral dissertation). University of South Florida.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*, 103–124.

Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology*, *42*, 1042–1050.

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94*, 265–292.

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology, 77*, 343–352.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*, 119–155.

Britton, J. C., Shin, L. M., Barrett, L. F., Rauch, S. L., & Wright, C. I. (2008). Amygdala and fusiform gyrus temporal dynamics: responses to negative facial expressions. *BMC Neuroscience*, *9*, 44.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the humanlikeness of digitally created faces. *Computers in Human Behavior*, *29*, 759–771.

Cacioppo, J. T., Martzke, J. S., Petty, R. E., & Tassinary, L. G. (1988). Specific forms of facial EMG response index emotions during an interview: From Darwin to the continuous flow hypothesis affect-laden information processing. *Journal of Personality and Social Psychology*, *54*, 592–604.

Cacioppo, J. T, Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*, 260–268.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*, 187–192.

Carver, C. S., & Scheier, M. F. (1978). Self-focusing effects of dispositional self-consciousness, mirror presence, and audience presence. *Journal of Personality and Social Psychology*, *36*, 324–332.

Chaminade, T., & Decety, J. (2001). A common framework for perception and action: Neuroimaging evidence. *Behavioral and Brain Sciences*, *24*, 879–882.

Chaminade, T., Zecca, M., Blakemore, S.-J., Takanishi, A., Frith, C. D., Micera, S., et al. (2010). Brain response to a humanoid robot in areas implicated in the perception of human emotional gestures. *Plos One*, *5*, e11577.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Chartrand, T. L., Maddux, W. W., & Lakin, J. L. (2005). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The New Unconscious* (pp. 334–361). Oxford: Oxford University Press.

Chin, M. G., Sims, V. K., DaPra, C. S., & Ballion, T. (2006). *Reliability of the Anthropomorphic Tendencies Scale.* Poster presented at Association for Psychological Science Convention. New York.

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental Social Psychology* (pp. 185–236). New York: Holt, Rinehart & Winston.

Dashiell, J. F. (1930). An experimental analysis of some group effects. *Journal of Abnormal and Social Psychology*, *25*, 190–199.

Dautenhahn, K. (1998). Elektronische Sisyphusse. Interview in *Pointer*. Retrieved July 27, 2005, from http://www.gmd.de/pointer/

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

Davidson, R., & Henderson, R. (2000). Electronic performance monitoring: a laboratory investigation of the influence of monitoring and difficulty on task performance, mood state, and self-reported stress levels. *Journal of Applied Social Psychology*, *30*, 906–920.

Davidson, R. J., & Hugdahl, K. (1995). *Brain Asymmetry*. Cambridge: MIT Press.

de Vries, M., Holland, R. W., Chenier, T., Starr, M. J., & Winkielman, P. (2010). Happiness cools the warm glow of familiarity: psychophysiological evidence that mood modulates the familiarity-affect link. *Psychological Science*, *21*, 321–328.

Dennett, D. C. (1989). *The Intentional Stance*. Cambridge: MIT Press.

Dennett, D. C. (1996). *Kinds of Minds: Towards an Understanding of Consciousness*. New York: Basic Books.

Dimberg, U. (1997). Facial reactions: Rapidly evoked emotional responses. *Journal of Psychophysiology*, *11*, 115–123.

Dimberg, U., & Lundquist, L. O. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology*, *30*, 151–159.

Dimberg, U., & Petterson, M. (2000). Facial reactions to happy and angry facial expressions: Evidence for right hemisphere dominance. *Psychophysiology*, *37*, 693–696.

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, *39*, 39–45.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*, 86–89.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: The design and perception of humanoid robot heads. *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, 321–326.

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience*, *6*, 90–97.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*, 177–190.

Duval, S., & Wicklund, R. A. (1972). *A Theory of Objective Self Awareness*. New York: Academic Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183–201.

Eger, E., Jedynak, A., Iwaki, T., & Skrandies, W. (2003). Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia*, *41*, 808–817.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Los Altos: Malor Books.

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating social connection through inferential reproduction. *Psychological Science*, *19*, 114–120.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864–886.

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When we need a human: Motivational determinants of anthropomorphism. *Social Cognition*, *26*, 143–155.

Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, *42*, 2213–2230.

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*, 724–731.

Eyssel, F., Hegel, F., Horstmann, G., & Wagner, C. (2010). Anthropomorphic inferences from emotional nonverbal cues: A case study. *Proceedings of the RO-MAN, 2010*, 646–651.

Eyssel, F., Kuchenbrandt, D., & Bobinger, S. (2011). Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism. *Proceedings of* *the 6th International Conference on Human-Robot Interaction,* 61–67.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology*, *60*, 229–240.

Fridlund, A. J., Sabini, J. P., Hedlund, L. E., Schaut, J. A., Shenker, J. I., & Knauer, M. J. (1990). Audience effects on solitary faces during imagery: Displaying to the people in your head. *Journal of Nonverbal Behavior*, *14*, 113–137.

Fridlund, A., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*, 567–589.

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 517–528.

Ghazanfar, A. A., & Turesson, H. K. (2008). How robots will teach us how the brain works. *Nature Neuroscience*, *11*, 3.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *Proceedings of the RO-MAN, 2003*, 55–60.

Gray, H. (1918). *Anatomy of the Human Body*. Philadelphia: Lea & Febiger.

Green, R. D., MacDorman, K. F., Ho, C.-C., & Vasudevan, S. (2008). Sensitivity to the proportions of faces that vary in humanlikeness. *Computers in Human Behavior*, *24*, 2456–2474.

Guérin, B. (1993). *Social Facilitation*. Cambridge: Cambridge University Press.

Hapgood, F. (2008, June). When Robots Live Among Us. *Discover Magazine*. Retrieved from http://discovermagazine.com/2008/jun/27-when-robots-live-among-us#.Un4GSH7QLY1

Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin*, *27*, 889–898.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, *2*, 96–99.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional Contagion*. Cambridge: Cambridge University Press.

Hatfield, E., Rapson, R. L., & Le, Y. (2011). Emotional contagion and empathy. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 19–30). Cambridge: MIT Press.

Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Hegel, F., Eyssel, F., & Wrede, B. (2010). The social robot “Flobi”: Key concepts of industrial design. *Proceedings of the RO-MAN, 2010*, 107–112.

Hess, U., Philippot, P., & Blairy, S. (1998). Facial reactions to emotional facial expressions: Affect or cognition? *Cognition and Emotion*, *12*, 509–531.

Hinds, P., Roberts, T., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, *19*, 151–181.

Hoffmann, L., Krämer, N. C., Lam-Chi, A., & Kopp, S. (2009). Media equation revisited: Do users show polite reactions towards an embodied agent? *Proceedings of the 9th International Conference on Intelligent Virtual Agents*, 159–165.

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, *12*, 183–195.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999).

Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

IFR Statistical Department (2013). World Robotics 2013 Service Robots. Retrieved from http://www.worldrobotics.org

Jones, E. E., & Gerard, H. B. (1967). *Foundations of Social Psychology*. Oxford: Wiley.

Kahn, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., et al. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, *48*, 303–314.

Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., et al. (2012). Do people hold a humanoid robot morally accountable for the harm it causes? *Proceedings of the 7th International Conference on Human-Robot Interaction*, 33–40.

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

Kanda, T., Ishiguro, H., Imai, M., & Ono, T. (2004). Development and evaluation of interactive humanoid robots. *Proceedings of the IEEE*, *92*, 1839–1850.

Kiesler, S., & Sproull, L. (1997). "Social" human-computer interaction. In B. Friedman (Ed.), *Human Values and the Design of Computer Technology* (pp. 191–199). Cambridge: Cambridge University Press.

Klamer, T., & Allouch, S. B. (2010). Zoomorphic robots used by elderly people at home. *Proceedings of the 27th International Conference on Human Factors in Computing Systems*.

Krämer, N. C. (2008). *Soziale Wirkungen virtueller Helfer: Gestaltung und Evaluation von Mensch-Computer-Interaktion*. Stuttgart: Kohlhammer.

Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2011). Minimal group - maximal effect? Evaluation and anthropomorphization of the humanoid robot NAO. In *Social Robotics* (pp. 104–113). Berlin, Heidelberg: Springer.

Kushnir, T., & Duncan, K. D. (1978). An analysis of social facilitation effects in terms of signal detection theory. *Psychological Record*, *28*, 535–541.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*, 261–273.

Langer, E. J. (1992). Matters of mind: Mindfulness/mindlessness in perspective. *Consciousness and Cognition*, *1*, 289–305.

Lanzetta, J. T., & Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *Journal of Personality and Social Psychology*, *56*, 543–554.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*, 776–785.

Likowski, K. U., Mühlberger, A., Seibt, B., Pauli, P., & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, *44*, 1065–1072.

Lundqvist, L. O. (1995). Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian Journal of Psychology*, *36*, 130–141.

Lütkebohle, I., Hegel, F., Schulz, S., Hackel, M., Wrede, B., Wachsmuth, S., & Sagerer, G. (2010). The bielefeld anthropomorphic robot head "Flobi." *Proceedings of the International Conference on Robotics and Automation, 2010*, 3384–3391.

Lynn, R. (1966). *Attention, Arousal and the Orientation Reaction*. London: Pergamon.

MacDorman, K. F., & Ishiguro, H. (2006a). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, *7*, 297–337.

MacDorman, K. F., & Ishiguro, H. (2006b). Opening pandoras box: Reply to commentaries on "The uncanny advantage of using androids in social and cognitive research." *Interaction Studies*, *7*, 361–368.

Mataric, M. J. (2008). Human-centered socially assistive robotics for rehabilitation: Is it time for a robot coach? *Proceedings of Virtual Rehabilitation, 2008*.

McHugo, G. J., Lanzetta, J. T., & Bush, L. K. (1991). The effect of attitudes on emotional reactions to expressive displays of political leaders. *Journal of Nonverbal Behavior*, *15*, 19–41.

McIntosh, D. N. (1996). Facial feedback hypotheses: Evidence, implications, and directions. *Motivation and Emotion*, *20*, 121–147.

McIntosh, D. N. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin*, *37*, 31–42.

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science*, *9*, 295–302.

Meltzoff, A. N., & Prinz, W. (Eds.). (2002). *The Imitative Mind.* Cambridge: Cambridge University Press.

Milgram, S. (1974). *Obedience to Authority: An Experimental View*. London: Tavistock.

Moody, E. J., & McIntosh, D. N. (2006). Mimicry and Autism: Bases and Consequences of Rapid, Automatic Matching Behavior. In S. J. Rogers & J. H. G. Williams (Eds.), *Imitation and the Social Mind: Autism and Typical Development* (pp. 71–95). New York: Guilford Press.

Moody, E. J., McIntosh, D. N., Mann, L. J., & Weisser, K. R. (2007). More than mere mimicry? The influence of emotion on rapid facial reactions to faces. *Emotion*, *7*, 447–457.

Moore, R. K. (2012). A bayesian explanation of the "uncanny valley" effect and related psychological phenomena. *Scientific Reports*, *2*, 864.

Mori, M. (1970). The uncanny valley. *Energy*, *7*, 33–35.

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, *121*, 47–57.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., et al. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *Journal of Neurophysiology*, *82*, 1610–1614.

Nass, C. I., Moon, Y., Morkes, J., Kim, E.-Y., & Fogg, B. J. (1997). Computers are social actors: A review of current research. In *Human Values and the Design of Computer Technology* (pp. 137–162). Stanford: Center for the Study of Language and Information.

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

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*, 669–678.

Nass, C., Isbister, K., & Lee, E. J. (2000). Truth is beauty: Researching embodied conversational agents. In *Embodied Conversational Agents* (pp. 374–402).

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the 1994 SIGCHI Conference on Human Factors in Computing Systems*, 72–78.

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 16598–16603.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review*, *9*, 184–211.

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, *7*, 437–454.

Norman, D. A. (1980). Twelve issues for cognitive science. *Cognitive Science*, *4*, 1–32.

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *12*, 481–494.

Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science*, *11*, 62–66.

Page, M. C., Braver, S. L., & MacKinnon, D. P. (2003). *Levine's Guide to SPSS for Analysis of Variance* (2nd ed.). London: Lawrence Erlbaum Associates, Inc.

Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors*, *49*, 1054–1060.

Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index. Retrieved from http://psydok.sulb.uni-saarland.de/volltexte/2009/2363/

Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. *Proceedings of the 2nd International Conference on Human-Robot Interaction*, 145–152.

Premack, D., & Premack, A. J. (1995). Origins of human social competence. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 205–218). Cambridge: MIT Press.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515–526.

Pütten, von der, A. M., Krämer, N. C., & Eimler, S. C. (2011). Living with a robot companion - Empirical study on the interaction with an artificial health advisor. *Proceedings of the 13th International Conference on Multimodal Interfaces*, 327–334.

Pütten, von der, A. M., Krämer, N. C., Gratch, J., & Kang, S. (2010). “It doesn’t matter what you are!” Explaining social effects of agents and avatars. *Computers in Human Behavior*, *26*, 1641–1650.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science*, *9*, 45–48.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: Cambridge University Press.

Rickenberg, R., & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. *Proceedings of the 2000 SIGCHI Conference on Human Factors in Computing Systems*, 49–56.

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin*, *95*, 52–77.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rosenthal-von der Pütten, A. M., Schulte, F. P., Eimler, S. C., Hoffmann, L., Sobieraj, S., Maderwald, S., et al. (2013). Neural correlates of empathy towards robots. *Proceedings of the 8th International Conference on Human-Robot Interaction*, 215–216.

Rosenthal-von der Pütten, A. M., Krämer, N. C., Hoffmann, L., Sobieraj, S., & Eimler, S. C. (2012). An experimental study on emotional reactions towards a robot. *International Journal of Social Robotics*, *5*, 17–34.

Sanders, G. S., & Baron, R. S. (1975). The motivating effects of distraction on task performance. *Journal of Personality and Social Psychology*, *32*, 956–963.

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience*, *7*, 413–422.

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*, 718–730.

Schmitt, B., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and social facilitation: One more time. *Journal of Experimental Social Psychology*, *22*, 242–248.

Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*, 433–443.

Spence, K. W. (1956). *Behavior Theory and Conditioning.* New Haven: Yale University Press.

Sprengelmeyer, R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *265*, 1927–1931.

Sproull, L., Subramani, M., Kiesler, S., Walker, J., & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, *11*, 97–124.

Sundar, S. S., & Nass, C. (2000). Source orientation in human-computer interaction. *Communication Research*, *27*, 683–703.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*, 149–178.

Tapus, A., Mataric, M., & Scassellati, B. (2007). Socially assistive robotics. *Robotics & Automation Magazine*, *14*, 35–42.

Thompson, E., & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, *5*, 418–425.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology, 9*, 507–533.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*, 585–594.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron*, *30*, 829–841.

Walker, J. H., Sproull, L., & Subramani, R. (1994). Using a human face in an interface. *Proceedings of the 1994 SIGCHI Conference on Human Factors in Computing Systems*, 85–91.

Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, *5*, 219–232.

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*, 410–435.

Weyers, P., Mühlberger, A., Kund, A., Hess, U., & Pauli, P. (2009). Modulation of facial reactions to avatar emotional faces by nonconscious competition priming. *Psychophysiology*, *46*, 328–335.

Williams, J. H. G., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews*, *25*, 287–295.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winkielman, P., Hall, L. O., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*, 799–806.

Winkielman, P., Schwarz, N., & Reber, R. (2000). Affective and cognitive consequences of visual fluency: When seeing is easy on the mind. *Visual Persuasion*.

Woods, S., Dautenhahn, K., & Kaouri, C. (2005). Is someone watching me? Consideration of social facilitation effects in human-robot interaction experiments. *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation*, 53–60.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274.

Zajonc, R. B. (1980). Compresence. In P. B. Paulus (Ed.), *Psychology of Group Influence* (pp. 35–60). Hillsdale: Lawrence Erlbaum Associates, Inc.

Zajonc, R. B., Heingartner, A., & Herman, E. M. (1969). Social enhancement and impairment of performance in the cockroach. *Journal of Personality and Social Psychology*, *13*, 83–92.

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2004). Effects of virtual human presence on task performance. *Proceedings of the 14th International Conference on Artificial Reality and Telexistence*, 174–181.

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2007). Social responses to virtual humans: Implications for future interface design. *Proceedings of the 2007 SIGCHI Conference on Human Factors in Computing Systems*, 1561–1570.

Declaration

I herewith declare after §8 of the Rahmenpromotionsordnung of Bielefeld University that I am in knowledge of the doctoral regulations of the faculty of psychology and sport science of Bielefeld University (Promotionsordnung der Fakultät für Psychologie und Sportwissenschaft vom 1.8.2011), that this thesis is my own work, that I have not received any undocumented help, and that I have used only the sources listed. Neither this thesis nor parts of this thesis have been accepted or are currently being submitted for any other degree or qualification at this university, outside of this faculty, or elsewhere. As yet, I have not undertaken any unsuccessful dissertation attempts at this university, outside of this faculty, or elsewhere.

Nina Riether Bielefeld, December 2013

Gedruckt auf alterungsbeständigem Papier, ISO 9706