Actions Speak Louder Than Assumptions

Empirical Investigations on the Perception of Artificial Entities Considering Expectations, Attributions, and Behavior

Von der Fakultät für Ingenieurwissenschaften,

Abteilung Informatik und Angewandte Kognitionswissenschaft

der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

genehmigte kumulative Dissertation

von

Aike Christiane Horstmann

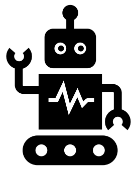
aus

Fulda, Deutschland

Gutachterin 1: Prof. Dr. Nicole Krämer

Gutachterin 2: Prof. Dr. Astrid Rosenthal-von der Pütten

Datum der mündlichen Prüfung: 16.02.2021



“We are fascinated with robots

because they are reflections of ourselves”

Ken Goldberg

ACKNOWLEDGEMENTS

“Are you sad because you're on your own?

No, I get by with a little help from my friends”

- John Lennon

On the adventurous journey resulting in this dissertation thesis, I was never on my own thanks to all those wonderful people who accompanied me. At this point I want to try to express the gratitude I feel for those people, although it is going to be hard to put this into a few words.

First, I want to thank my supervisor Prof. Dr. Nicole Krämer for all her support, confidence, trust, encouragement, feedback, and little pushes whenever they were needed. I deeply appreciate the freedom you gave me to develop my own research ideas and study plans while always being available with valuable feedback and constructive criticism. I also want to thank Prof. Dr. Astrid Rosenthal-von der Pütten for being such an inspiring robot research pioneer and for accepting to be the second examiner for my thesis.

Next, I thank the wonderful SP team for all the conferences, milestone celebrations, feedback and motivation, and just in general for going through all the ups and downs of academic life together. I have so many marvellous memories of the last years which I am going to cherish forever. Particularly, I am thankful to my roommate Jessica Szczuka for all the fun times as well as deep talks. Furthermore, I thank Yannic Meier, Evgenia Princi, Judith Meinert, Filipa Stoyanova, Jan Philipp Kluck, Marcel Finkel, and Lea Frentzel-Beyme who gave me the emotional as well as professional support to get through these last tough months. Likewise, I want to thank my friends outside of SP, who were by my side not only but particularly the last years.

I am so grateful for my amazing family – the A-Team including Achim, Ann kathrin, Arne, and Antje Horstmann. Thank you for always having my back, for all the love and support that made me able to get where I am and more importantly made me the person I am today.

Last and foremost I want to express my deepest gratitude to my partner in crime and in everything else, Ercan Karaca, who carried and sometimes pushed me whenever I needed it. I thank you so much for all you have done, for taking care of me when I could not, for having my back, for celebrating and mourning with me, for encouraging and always believing in me. Sen ve ben. Birlikte geleceğe. Seni seviyorum, aslanım.

ANNOTATION OF THE PAPERS INCLUDED IN THE CUMULUS

Research Paper 1:

Horstmann, A. C., & Krämer, N. C. (2019). Great Expectations? Relation of Previous Experiences With Social Robots in Real Life or in the Media and Expectancies Based on Qualitative and Quantitative Assessment. *Frontiers in Psychology, 10*, 939. https://doi.org/10.3389/fpsyg.2019.00939

Research Paper 2:

Horstmann, A. C., & Krämer, N. C. (2020). Expectations vs. Actual Behavior of a Social Robot: An Experimental Investigation of the Effects of a Social Robot's Interaction Skill Level and Its Expected Future Role on People's Evaluations. *PLoS ONE, 15*(8), e0238133. https://doi.org/10.1371/journal.pone.0238133

Research Paper 3:

Horstmann, A. C., & Krämer, N. C. (2020). The Fundamental Attribution Error in Human-Robot Interaction: An Experimental Investigation on Attributing Responsibility to a Social Robot for Its Pre-Programmed Behavior. *Manuscript submitted for publication in the International Journal of Social Robotics.*

Research Paper 4:

Horstmann, A. C., Gratch, J. & Krämer, N. C. (2020). I Just Wanna Blame Somebody, Not Something! Reactions to a Computer Agent Giving Negative Feedback Based on the Instructions of a Person. *Manuscript under review for publication in the International Journal of Human-Computer Studies.*

ZUSAMMENFASSUNG

Die Technologie von artifiziellen Entitäten entwickelt sich zurzeit rasant weiter, was zur Folge hat, dass diese vermehrt in den verschiedensten privaten sowie öffentlichen Bereichen zum Einsatz kommen. Da artifizielle Entitäten technologischer Natur sind, dennoch aber Lebendigkeit durch ihr Aussehen und Verhalten suggerieren, ist es schwierig, sie eindeutig zu klassifizieren. Dies wiederum hat verschiedene Auswirkungen auf ihre menschlichen Interaktionspartner und löst verschiedene Reaktionen aus, welche noch nicht vollständig erforscht wurden. Um erfolgreiche sowie angenehme zukünftige Interaktionen mit artifiziellen Entitäten zu gestalten, ist es wichtig genauer zu untersuchen, welche Faktoren die Wahrnehmung beeinflussen und wie diese Faktoren miteinander interagieren. Dies wird durch die Studien des vorliegenden Kumulus thematisiert, welche Folgendes untersuchen: a) Erwartungen zusammen mit technologischer Expertise als Wahrnehmungsfilter der Realität, b) Attributionsprozesse in Interaktionen mit künstlichen Entitäten und c) die Wirkung des Verhaltens von artifiziellen Entitäten auf die Gesamtwahrnehmung.

Im ersten Schritt lag der Schwerpunkt von Studie I und II (Forschungsarbeit I) darin, die Erwartungen der Menschen bezüglich künstlicher Entitäten und die Entstehung dieser Erwartungen zu erforschen. Dazu wurden persönliche Erfahrungen, Medienrezeption und technologische Expertise als Einflussfaktoren betrachtet. Insbesondere eine erhöhte Medienrezeption, welche auch die Rezeption von Science-Fiction-Formaten umfasste, führte zu hohen Erwartungen bezüglich der Fähigkeiten sowie des generellen zukünftigen Einsatzes künstlicher Entitäten. Technologische Expertise scheint zudem Ängste im Zusammenhang mit künstlichen Entitäten zu reduzieren, während die Kenntnis von "bösartigen" artifiziellen Science-Fiction-Charakteren negative Erwartungen verstärkt.

Auf Basis der beiden in den Massenmedien prominentesten Ansichten bezüglich artifizieller Entitäten, untersuchte Studie III (Forschungsarbeit II) wie ein sozialer Roboter wahrgenommen wird, welcher entweder mit der Absicht ein hilfreicher Assistent oder ein bedrohlicher Konkurrent zu werden dargestellt wird. Zudem präsentierte der Roboter entweder dürftige oder hochentwickelte Interaktionsfähigkeiten. Der Assistenz-Roboter wurde als sozial umgänglicher empfunden im Vergleich zu dem Konkurrenz-Roboter. Das Niveau der Interaktionsfähigkeiten hatte jedoch mit Abstand den stärksten Einfluss auf die Gesamtwahrnehmung des Roboters.

Studie IV (Forschungsarbeit III) untersuchte, wie die Wahrnehmung eines Roboters, welcher entweder positives oder negatives Feedback gibt, dadurch beeinflusst wird, ob sein Handeln als autonom oder streng vorprogrammiert dargestellt wird. Mehr Agency, Verantwortlichkeit und

Kompetenz wurden dem autonom agierenden Roboter zugeschrieben, während die Valenz des Feedbacks den größten Einfluss auf die Gesamtwahrnehmung des Roboters hatte.

In Studie V (Forschungsarbeit IV) wurde darüber hinaus untersucht, wie negatives Verhalten einer artifiziellen Entität wahrgenommen wird, wenn sie entweder von einer Person, einem Computer oder einer Mischung aus Person und Computer gesteuert wird und zudem entweder über hohe oder geringe Expertise verfügt. Die Ergebnisse zeigten nur geringe Unterschiede, die darauf hindeuten, dass Agency aber auch Schuldvorwurf und negativer Affekt bei der Steuerung durch eine Person am höchsten waren und am niedrigsten, wenn ein Computerprogramm die Handlungen bestimmte, während sich die Mischform irgendwo dazwischen bewegte.

Zusammenfassend wurden anhand der fünf Studien die individuellen sowie kombinierten Einflüsse von Erwartungen, Attributionsprozessen und Verhaltensweisen auf die Wahrnehmung künstlicher Entitäten eingehend untersucht. Erwartungen sowie die technologische Expertise der Menschen bildeten einen Rahmen für die Wahrnehmung. Insbesondere die am häufigsten in den Medien vertretenen Erwartungen – Assistenz versus Konkurrenz – beeinflussten die wahrgenommene soziale Umgänglichkeit artifizieller Entitäten. Was den Einfluss von Attributionen betrifft, so verstärkte die Darstellung einer artifiziellen Entität als autonom agierend ihre wahrgenommene Agency, Verantwortlichkeit und Kompetenz. Abgesehen von diesen Faktoren, die sich auf die Wahrnehmung bestimmter Merkmale einer artifiziellen Entität beschränken, stellt das Verhalten von artifiziellen Entitäten den größten Einflussfaktor auf die Gesamtwahrnehmung dar. Genauer gesagt führt als angenehm empfundenes Verhalten dazu, dass die artifizielle Entität in einem positiven Licht wahrgenommen wird, während abstoßend empfundenes Verhalten zu einer negativen Wahrnehmung führt. Um erfolgreiche und erstrebenswerte Interaktionen mit artifiziellen Entitäten zu gestalten, sollten auf Basis der vorliegenden Ergebnisse die folgenden Aspekte beachtet werden: a) wünschenswerte Erwartungen und eine hohe technologische Expertise können Ängste abbauen und die wahrgenommene soziale Umgänglichkeit verbessern, b) wahrgenommene Autonomie fördert die Zuschreibung von Agency, Kompetenz und Verantwortlichkeit und c) eine sorgfältige Gestaltung von positiv empfundenen Verhalten führt zu einer positiven Gesamtbewertung der artifiziellen Entität und der Interaktion mit dieser.

ABSTRACT

The technology behind artificial entities is rapidly advancing which entails that artificial entities are progressively used in various areas ranging from people's homes to public institutions. Since artificial entities are technological in nature but suggest liveliness via their appearance and behavior, it is challenging to classify them without ambiguity. This in turn has various effects on and elicits different reactions of people interacting with them, which are not fully explored yet. In order to design successful and pleasant future interactions with artificial entities, it is crucial to examine more thoroughly which factors have what kind of influence on how they are perceived and how these factors interact with each other. This is addressed by the studies of the cumulus which particularly investigate a) expectations alongside with technological expertise as perceptual filters of the reality, b) attribution processes in interactions with artificial entities, and c) the effects of artificial entities' behaviors on overall perceptions.

As a first step, the focus of study I and II (research paper I) was to explore people's expectations regarding artificial entities and how they develop by considering personal experiences, reception of mass media, and technological expertise as influencing factors. Particularly mass media reception including science fiction formats led to enhanced expectations regarding the skills and general future role of artificial entities. Technological expertise appeared to have a diminishing effect on fears related to artificial entities, while recalling "bad" artificial science fiction characters enhanced negative expectancies.

Drawing on the two views on artificial entities which are most prominently represented in mass media, study III (research paper II) examined the effect of portraying a social robot as aiming to become a helpful assistant or threatening competitor in combination with having the robot display either poor or elaborated interaction skills. The assistant-expectation caused people to perceive the robot as more sociable than the competitor-expectation, however, the level of interaction skills was able to determine the overall perception with large effect sizes.

Study IV (research paper III) looked at how portraying a social robot as acting autonomously or being strictly pre-programmed influences how it is perceived in combination with having the robot provide either positive or negative feedback. More agency, responsibility, and competence were attributed to the autonomously acting robot, while the valence of the feedback had the greatest impact on the overall perception.

In study V (research paper IV) it was further examined how having a person, computer, or a mix form of the two control an agent's behavior combined with high or low expertise influences how its offensive behavior is perceived. Only minor differences could be found which indicate

that agency, blame, and negative affect were highest with a person and lowest with a computer program controlling the agent's action, while the mix form settles somewhere between.

Summing up, by means of these five studies, the individual and combined influences of expectations, attributions, and behaviors on the perception of artificial entities were explored in depth. Expectations and people's technological expertise were found to form a perceptual frame. Particularly, expectations regarding the two most common views on artificial entities in mass media – assistant versus competitor – have an effect on the perceived sociability of artificial entities. Regarding the influence of attributions, presenting an artificial entity as acting autonomously enhances its perceived agency, responsibility, and competence. Besides these factors having confined effects on the perception of certain characteristics of an artificial entity, its behavior constitutes the greatest factor influencing the overall perception. More specifically, pleasant behavior leads to perceiving the artificial entity in a rather favorable light, while repulsive behavior leads to a negative perception. To foster successful and desirable interactions with artificial entities, based on the present results the following should be kept in mind: a) desirable expectations and high technological expertise reduce fears and enhance the perceived sociability, b) autonomy promotes the attribution of agency, competence, and responsibility, and c) a careful design of positively perceived behavior leads to a positive overall evaluation of the artificial entity and the interaction with it.

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

Artificial entities, whether referring to robots with real-world physical embodiments or virtual agents presented on a screen, fascinate all kinds of people across the world and inspired the works of many science fiction authors over the years (Asimov, 1947, 1950; Clarke, 1968; Dick, 1968). As a consequence, most people grew up with the beloved R2-D2 from Star Wars as well as with the threatening Terminator. However, what long appeared in science fiction only, now begins to conquer the real world. With the rapid advancement of the technology behind artificial social interaction partners, people are now confronted with the prospect of encountering these artificial entities on a daily basis in their private as well as work-related environments. This could be interacting with a chat bot when contacting a company's customer service, using a speech assistant like Amazon's Alexa to play music or command smart home devices, or collaborating with a social robot when taking care of elderlies or children. Since artificial entities are clearly technological in nature but create the impression to be alive (Kahn et al., 2011), questions concerning people's perception of these entities arise. It is particularly pivotal to examine factors and mechanisms influencing people's perceptions of artificial entities in depth since they will ultimately determine the acceptance, usage intentions, as well as actual usage of these kinds of technologies (Davis, 1989, 1993). Beyond that, learning how people form perceptions regarding artificial entities will teach us more about how humans feel, think, and behave when interacting with non-human interaction partners. Ultimately, this will also extend our knowledge regarding human perception mechanisms in general. By means of social interactions with artificial entities, we are able to systematically examine social mechanisms and to extract specific perceptional and behavioral responses to certain factors while controlling for others. This is particularly feasible by employing artificial compared to human interaction partners, since artificial entities are more controllable and able to systematically simulate certain expressions. Against this background, the aim of this thesis is to examine expectations, attributions, as well as behaviors and how they influence people's perception of their interaction partners, particularly in the context of interactions with artificial entities.

For a start, expectations are believed to function as perceptual filters of the reality (Burgoon, 1993; Burgoon & Le Poire, 1993) and thus may shape how artificial entities are perceived considerably. Expectations can be formed based on a) categories associated with typical behaviors, attitudes, and emotional reactions, b) personal experiences as direct observations of behavior, and c) reputation based on reports of experiences by others (Andersen & Klatzky, 1987; Darley & Fazio, 1980). Regarding the assignment of a category, artificial entities pose a great and novel challenge. As mentioned before, artificial entities are clearly of non-human and technological nature, however, they possess characteristics of a living being and several studies show that people generally tend to react to artificial entities as if they were interacting with human interaction partners (Nass & Moon, 2000; Reeves & Nass, 1996). These circumstances make it hard for people to assign a clear category to them (Kahn et al., 2011; Severson & Carlson, 2010), which is described as uncertainty at category boundaries (Ramey, 2006). Since most people also lack personal experiences with sophisticated artificial entities, they likely rely for most parts on information by third party sources (Banks, 2020). Because mass media are wide-spread and easily accessible by a great audience (Nisbet et al., 2002), they are believed to be important means by which people form their expectations (Banks, 2020). These expectations in turn are often inflated and may not reflect actual state-of-the-art technology, which leads to a gap between people's expectations and the reality of artificial entities and causes biased perceptions (Komatsu & Yamada, 2011; Kwon et al., 2016). Thus, this work takes into account people's expectations as important aspect influencing the perception of artificial entities.

Additionally, since artificial entities are technological devices or applications, people's technological expertise may play a considerable role for the perception of these entities. Here, people's attitude towards technologies as well as their locus of control when handling technological devices should have an impact. Technical affinity, which reflects the extent people are attracted to technologies, should have a consistently positive influence on people's perception of artificial entities since they represent a particularly fascinating new technology (Karrer et al., 2009; Morahan-Martin & Schumacher, 2007). Individuals with a high locus of control when handling technology, however, may be able to evaluate artificial entities and their abilities more realistically (Beier, 1999).

The next influencing factor for the perception of artificial entities considered in this thesis is the attribution of the causality of and intentionality behind an artificial entity's behavior. As the attribution theory by Kelley (1973) suggests, people always seek to identify whether a behavior is caused by internal or external factors. This is extended by more recent research which emphasizes the importance of the motives which are believed to have guided a person's behavior for how this person is perceived and evaluated (Malle et al., 2001; Reeder et al., 2002). In order to ascribe causality as well as intentionality to an artificial entity, agency is considered as prerequisite. Agency can be described as the capability to perform actions (Himma, 2009), independently and free of choice (Barker, 2005) and is one of the most discussed aspects in the field of interactive devices. Since one of the main purpose of artificial entities is to interact with humans and to take over traditionally human tasks, the attribution of moral agency in particular moves into the center of attention (Banks, 2019; Sullins, 2006). Moral agency comprises morality and agency and is described as “intrinsic potential for enacting moral functions” (Banks, 2019; p. 363). Potential prerequisites for the perception of moral agency in artificial entities are autonomy, intentionality, and responsibility (Sullins, 2006). Artificial entities may not be able to fulfill these qualifications by definition due to a lack of human characteristics such as personal emotions, intentions, beliefs, and desires. However, since people appear to perceive artificial entities as somehow alive as reaction to certain social cues (Reeves & Nass, 1996), the focus should be put on the perception of agency while neglecting the question whether artificial entities are able to possess agency from a logical point of view. Based on previous research, how actors are perceived depends on whether their actions are attributed internally to the actor’s disposition or externally to situational or other external forces (Kelley, 1973; Kelley & Michela, 1980). Furthermore, whether the actions are perceived to be executed deliberately and which motives are perceived to have guided those actions will also have an impact on the perception of the actor (Malle et al., 2001; Reeder et al., 2002). These processes of attributing causality as well as intentionality should be affected by the extent to which artificial entities are perceived to be acting autonomously, intentionally, and with full responsibility (Kelley, 1973; Malle et al., 2001; Reeder et al., 2002; Sullins, 2006). In turn, this should influence how the entity’s actions are interpreted and consequently how it is perceived in general.

As last substantial aspect considered for this cumulus, the behavior of artificial entities is believed to have a substantial impact on how they are perceived. Behavior includes several components such as gaze, body movements, gestures, postures, facial expressions, and voice, which may be used to express emotions, adhere to social conversation norms, send natural cues and for sophisticated dialogue management (Cassell & Thórisson, 1999; Häring et al., 2011; Looije et al., 2010). These behaviors in their diverse forms can have various effects on people’s perception, individually and in combination with each other. For instance, after a robot behaved in a rude manner, people reported dislike and increased their physiological as well psychological distance to it as reaction to direct gaze by their non-human interaction partner, which did not happen when it behaved in a polite, likeable way before (Mumm & Mutlu, 2011). In a different study, after behaving in a functional, machine-like way people hesitated longer to switch off an emotionally protesting robot, which was less the case with a socially behaving robot (Horstmann et al., 2018). In general, behavior has not an inherently positive or negative influence on the perception of artificial entities, the valence depends on the content of the behavior (Rickenberg & Reeves, 2000). Essentially, behavior increases the social presence of artificial entities, which results in interpreting their behavior within a social context. As a consequence, an artificial entity's behavior determines or at least heavily influences how it is perceived during an interaction as well as evaluated after (Rickenberg & Reeves, 2000).

Accordingly, behavior in addition to people's expectations and technological expertise as well as attributions of causality and intentionality need to be taken into account as important influencing factors for the perception of artificial entities. The individual as well as combined influences of these factors are the central research objects of this cumulus. The findings of the studies, which form the centerpiece of this thesis, shall contribute to a deepened understanding of the perception of artificial entities, particularly how these perceptions emerge and how they may shape future interactions between humans and artificial entities.

The present work begins with a collection of central theories from social and media psychology to draw conclusions and transfer insights from human-human to human-machine interaction as well as to derive the research objectives for this thesis. This is accompanied by relevant empirical studies investigating the formation of perceptions in general and regarding artificial entities in particular, with focus on expectations and technological expertise as well as attribution processes and behavior. This is followed by a summary of the five studies included in the cumulus, which tackle the questions regarding the influences of different factors on the perception of artificial entities from distinct perspectives employing a selection of different methodological approaches. This work is completed by a thorough discussion of the outcomes of the studies, including limitations and prospects for future research as well as concrete theoretical and practical implications based on the presented findings.

2. THEORETICAL BACKGROUND

2.1. Artificial Entities – A Brief Overview

The term artificial entity describes non-human interaction partners and comprises several technologies such as domestic, service, industrial, and social robots, virtual computer agents, chat bots, voice assistants, and automated phone call or negotiation systems. An artificial entity is defined as man-made and with real existence in this world (Sugiyama, 2019), which means it was created by humans to exist in and interact with the real world. With the rapid advancement of these technologies, they become increasingly omnipresent in people's everyday lives. There are several kinds of artificial entities with which many people already interact with on a daily basis. For instance, a survey showed that the majority has used a voice assistant, such as Alexa or Siri, before and more than half reported to use this technology on a daily basis (PwC, 2018). With regard to virtual assistants and chatbots, a majority of business leaders already use them or plan on using them in the next couple of years (Oracle, 2016). Likewise, the area of personal service robots has a high expected growth rate (Visti, 2018). This may be enhanced by the fact that various forms of social robots are being prepared to be employed in several relevant contexts such as healthcare (Breazeal, 2011), elderly care (Broekens et al., 2009), support for autistic people (E. S. Kim et al., 2013) as well as in the service and entertainment sector for example as receptionists (Gockley et al., 2005), museum tour guides (Burgard et al., 1999), or sexual partners (Szczuka & Krämer, 2017).

For the present work, social robots as well as virtual agents are used in the studies and summarized under the term artificial entities. Virtual agents are defined as "virtual - often anthropomorphic - representations of computer interfaces that a user can socially identify with" (Burgoon et al., 2016; p. 24). A social robot is described by Bartneck and Forlizzi (2004) as "an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioral norms expected by the people with whom the robot is intended to interact" (p. 592). Both usually have a human-like or animal-like appearance and are able to display social means of expression (Nuñez & Rosenthal-von der Pütten, 2018). Despite slight preferences for social robots (depending on the task and environment), virtual agents and robots essentially elicit similar reactions in humans (Fridin & Belokopytov, 2014; Holz et al., 2009; Krämer et al., 2012; J. Li et al., 2016; Rosenthal-von der Pütten et al., 2016). Although virtual agents might be more limited in their physical expressiveness due to their confinement to a screen (Holz et al., 2009), they clearly meet the standards for eliciting social reactions by Reeves and Nass (1996): interactivity, natural language, and filling of a traditionally human role. Moreover, virtual agents are usually also embodied and able to use this body for expressiveness in a similar way robots are able to (Holz et al., 2009). Based on these deliberations, social robots as well as virtual agents are included as research objects in this thesis and referred to as artificial entities.

Concerning the appearance of artificial entities, how their exterior is designed is highly relevant for people, for instance, because abilities and functionalities are derived from the presence of certain appearance characteristics (Goetz et al., 2003; Rosenthal-von der Pütten & Krämer, 2015). Artificial entities are not necessary designed to be life-like, realistic replications of humans, partly because too realistic representations were observed to elicit negative reactions such as eeriness and revulsion (described as uncanny valley phenomenon; Mori, 1970;

Rosenthal-von der Pütten & Krämer, 2015; Seyama & Nagayama, 2007). However, due to their fields of application, multi-facet social interactions often constitute a pivotal component (Young et al., 2011). Consequently, artificial entities are usually equipped with some humanlike features in their appearance and/or behavior to make those interactions more natural and thus intuitive for the users (Fong et al., 2003; Huber et al., 2008; for an overview of different types of social robots see Leite et al., 2013). For instance, the robot Kismet was designed to express four types of social responses (affective, exploratory, protective, and regulatory) in order to convey intentionality and facilitate social interactions (Breazeal & Scassellati, 1999). When an artificial entity is able to convey internal states, this allows their human interaction partner to predict their behavior, respond to social approaches, and engage in natural communication with them (Breazeal & Scassellati, 1999). Since behavior is believed to have a greater influence than appearance alone (Rickenberg & Reeves, 2000), this work focuses on the influence of behavioral aspects. As reaction to the usage of particularly social behavior cues, people often treat interactive devices as if they were real persons, which is why an extensive body of research focused on comparing people's perceptions and reactions in interactions with humans to interactions with artificial entities (Nass & Moon, 2000; Reeves & Nass, 1996).

2.2. Perceptions of Humans as Research Basis

Since people often treat artificial entities similar to how they treat other people (Reeves & Nass, 1996), it makes sense to first look at in which regards perceptions of humans resemble and also how they differ from perceptions of artificial entities. Previous social psychological research focused on human experience and behavior in context of social interactions that only took place between humans. Now that artificial entities appear on the scene, social interactions also take place with artificial entities and insights from human-human interactions potentially also apply in these contexts. Some of the most relevant findings regarding factors influencing the human perception of other humans are presented in the following. At the end of this section, it is further outlined how these findings can be transferred to interactions with artificial entities based on previous research in this field.

2.2.1. Expectations Influencing the Perception of Humans

As described by the uncertainty reduction theory (Berger & Calabrese, 1975), during an initial encounter with an unfamiliar interaction partner, one of people's primary concern is to reduce the uncertainty which derives from lacking information about the other. For this purpose, people attempt to gather enough information to be able to predict and explain the other's behavior

(Berger & Calabrese, 1975). In doing so, first impressions are formed automatically, i.e., fast and without much effort (Asch, 1946). These impressions are often affected by the assignment of a social category which are based on trait concepts or social stereotypes (Andersen & Klatzky, 1987). Social categories help to form impressions because they are believed to have predictive power over their members' typical behaviors, emotional reactions, personality traits, attitudes, and values (Andersen & Klatzky, 1987). Accordingly, due to people's need to reduce uncertainty (Berger & Calabrese, 1975), expectations are formed based on general social norms, the ascription of a social category, and/or personal knowledge about the other (Burgoon & Hale, 1988; Darley & Fazio, 1980). Expectations in turn shape and filter people's perceptions by ascribing different meanings to the same kind of behavior (Burgoon, 1993; Burgoon & Le Poire, 1993). For instance, increased proximity during a conversation may be perceived as pleasant with a friend, but as unpleasant with a stranger (Burgoon & Jones, 1976). With friends, proximity as sign of affiliation and familiarity is expected, which is not the case with strangers (Burgoon & Jones, 1976; Burgoon & Walther, 1990). Thus, it can be concluded that expectations can influence how a person and their behavior are perceived. Another aspect that also appeared to have a great impact on how people are perceived are attribution processes which in the end lead to perceiving a person responsible for its actions or not.

2.2.2. Attribution Processes Influencing the Perception of Humans

Calling back to mind the uncertainty reduction theory, we learned that there is a fundamental need to reduce uncertainty in social interactions by gathering information to understand and predict the other's behavior (Berger & Calabrese, 1975). Extending this, the attribution theory describes how people always attempt to figure out which factors cause or contribute to other persons' behaviors, likewise with the aim to understand and predict future behaviors (Kelley, 1973; Kelley & Michela, 1980; Ross, 1977). In general, causes are divided into internal and external causes. In case of the first, behavior is attributed to a person's disposition and in case of the latter, behavior is attributed to situational or contextual factors (Kelley, 1973).

Following Kelley's covariation model, there are mainly three factors which determine whether an action should be attributed to an internal or an external cause (Kelley, 1973). The first factor is called consensus and describes that a person's response reflects other persons' responses to the same stimulus. The second factor, consistency, is high if the person of interest repeatedly responds in a similar way to a specific stimulus, over time as well as in different contexts. The third factor was named distinctiveness and describes whether the person exclusively responds this way to this one stimulus or to other stimuli as well (Kelley, 1973). The person is held responsible for their behavior when consensus and distinctiveness are low and consistency high (Hewstone & Jaspars, 1987). The stimulus is seen as causing a person's reactions when consensus, consistency, and distinctiveness are high and situational factors are seen as primary source of cause when consensus is high, but distinctiveness and consistency are low (Hewstone & Jaspars, 1987).

An extensive body of researchers further suggests to extend the assumptions of the attribution theory, which focuses on the determination of causality, by considering the (perceived) motives eliciting an action and thus taking into account the intentionality behind a person's behavior (Malle et al., 2001; Malle & Hodges, 2007; Reeder et al., 2002; Reeder, 2009). According to them, different attributed motives can further affect how an act and the respective actor are perceived. For instance, in a study by Reeder et al. (2002), people perceived aggressive behavior to be selfishly motivated when there was a situational reward while aggression in response to situational provocation was perceived to be motivated by revenge and self-defense. The behavior perceived to be selfishly motivated was evaluated more negatively than the behavior perceived to be driven by revenge. Although there are external, situational factors present in both cases (either a reward or a provocation), the attributed motives eventually appeared to determine how the behavior was perceived. Accordingly, the authors conclude that "perceived motives play an important role in dispositional inference and pose a problem for models that focus primarily on perceived causality" (p. 789). Thus, in addition to determining the causality, these researchers further argue that it is pivotal to include the perception of motives for a certain behavior to understand how people make judgements about others' dispositions. To evaluate actions and those who executed those actions, people make attributions by considering the cause of and also the intentionality behind those actions (Malle et al., 2001).

All these attribution processes should be influenced by the degree to which a person is perceived to possess agency in the situation. There is no clear definition for agency, however, it may be described as the capacity to perform actions independently and voluntarily (Barker, 2005; Himma, 2009). Moral agency takes this one step further by determining whether an agent is able to act based on moral judgements, i.e., a notion of right or wrong (Bandura, 2002; Banks, 2019). As prerequisites for moral agency autonomy, intentionality, and responsibility are named (Sullins, 2006). In the end, the extent to which these aspects are perceived to be present will determine whether a person or external factors are held accountable for this person's actions (Kelley, 1973).

Let us take a closer look at these prerequisites. First, a person needs to be perceived to be acting with autonomy, i.e., on its own and not coerced by external forces (Kelley, 1973). As Winner (1977) phrases it: “to be autonomous is to be self-governing, independent, not ruled by an external law of force” (p. 16). This independence may be cognitive such as making one's own decisions and/or physical such as performing one's own actions (Banks, 2019). In the end, agency is only attributed to a person who is perceived to be acting autonomously, regardless of whether this is actually the case or not. Second, the extent a person is perceived to act intentionally will further decide whether the person is perceived to possess agency and to be accountable for their actions (Sullins, 2006). Intention stands for the planning of, commitment to, and control of behaviors or actions to satisfy certain needs (Fishbein & Ajzen, 2010; Kiss, 1992). Thus, intentionality describes whether an action was carried out by accident or with the purpose to achieve a certain goal. Since attributed motives may ultimately determine how an action and the respective actor are perceived, a closer look should be taken at the attribution of intentionality behind a certain behavior (Malle et al., 2001; Reeder et al., 2002). According to Jones and Davis (1965) “the attribution of intentions is that actions are informative to the extent that they have emerged out of a context of choice and reflect a selection of one among plural alternatives” (p. 219). Thus, a person’s actions and subsequent consequences are rather attributed to the person and influence how this person is perceived when they are believed to be deliberate and calculated (Sullins, 2006). Third, responsibility results from a person acting autonomously and intentionally. According to Sullins (2006), responsibility may result from contextual or role specific factors. For instance, teachers are responsible in their classrooms and nurses while taking care of their patients. These roles make people responsible to act morally.

Whether a person is perceived to be acting autonomously and with intentions will influence the extent to which the person is perceived to possess agency. This, in turn, will determine whether the cause of the person’s behavior is attributed to the person or situational factors (Kelley, 1973; Kelley & Michela, 1980; Sullins, 2006) as well as which motives are attributed to the person, which will then ultimately affect how this person is perceived in general (Malle et al., 2001; Reeder et al., 2002). This attribution of causality accompanied by other factors such as motives and the person’s role will determine whether the person is perceived responsible (Shaver, 1985; Sullins, 2006). As Shaver (1985) argues, the determination of causality is dichotomous and involves human agency, while responsibility is variable and a judgement based on several different dimensions.

Coming to a conclusion, many attribution processes take place in interpersonal settings because people strive to understand others' behaviors to reduce uncertainty and enhance predictability (Berger & Calabrese, 1975; Kelley, 1973). Here, actions that are perceived to be performed autonomously and intentionally convey a strong sense of agency and are attributed to the performer, which will then determine in addition with other factors how responsible this person is perceived (Kelley, 1973; Shaver, 1985; Sullins, 2006). The extent to which behaviors and their consequences are attributed internally to a person and their motives should ultimately have a great impact on how this person and their behaviors are perceived, which leads us to the role of the behavior itself. The perception of behavior may be affected by expectations as well as attribution processes as outlined before, however, the behavior and its meaning need to be looked at and considered individually as well.

2.2.3. Behavior Influencing the Perception of Humans

Previous research underlines the powerful and overarching effect of people's interactional behavior on how their actions are interpreted as well as how these people are perceived in general (Burgoon, 1993; Burgoon & Le Poire, 1993). Particularly when common interactional norms are violated, previous expectations are neglected and the action itself comes to the fore by having the strongest effects on communication outcomes (Burgoon & Le Poire, 1993). Likewise, the attribution of causality does not always appear to be executed as rational as described by Kelley's covariation model (Kelley, 1973). On the contrary, external factors are often neglected, while internal factors are overestimated (Ross, 1977). In other words, the focus lays on the influence of a person's disposition, while underestimating external forces such as the situational context, which was labelled as fundamental attribution error or overattribution effect (Jones & Nisbett, 1972; Nisbett & Ross, 1983; Ross, 1977). For instance, a study by Jones and Harris (1967) delivered evidence for this phenomenon with regard to chance-directed behaviors. A coin toss determined the attitude of an essay people were asked to write (pro or contra). Despite receiving information about how the attitude was assigned by chance, readers of the essays were still found to attribute the imposed attitude to the writers (Jones & Harris, 1967). Thus, the behavior was interpreted directly, without considering external factors contributing to it.

This appears to be further evidence for the overshadowing and direct influence of a person's behavior, particularly if it has a clear meaning, on people's perception and evaluation of this person. A behavior with a clear, unequivocal meaning is for example the display of an insulting gesture, while other behaviors such as proximity and distance hold multiple interpretation alternatives depending on contextual, personal, and relational aspects (Burgoon & Hale, 1988). In case of ambiguity, expectations and attributions seem to play a greater role in the interpretation and evaluation of behavior, which then affects how the person displaying the behavior is perceived (Burgoon, 1993; Kelley, 1973). Against this background, the question arises how these insights from human-human interactions can be used to receive a deepened understanding of the processes in human-machine interactions, particularly whether perceptions of artificial entities emerge and are influenced in a similar way as perceptions of other humans.

2.3. Transferring Insights from Perceptions of Humans to Perceptions of Artificial Entities

Interactions with artificial entities such as virtual agents or robots are increasingly designed to facilitate elaborate social interactions, for instance, by means of speech output and personable characters (Parise et al., 1999). Consequently, users' perceptions need to be understood and assessed by considering a social interaction framework rather than a restricted machine interaction framework (Parise et al., 1999). As a result, social psychological knowledge and methods are increasingly incorporated in human-machine research and differences as well as similarities between interactions with a human or artificial interaction partner are explored with a rising effort (Krämer et al., 2012).

According to Reeves and Nass (1996), "individuals' interactions with computers, television, and new media are fundamentally social and natural, just like interactions in real life" (p. 5). In other words, when people interact with media which fulfill certain criteria – a few fundamental social cues, like interactivity, natural speech, and filling a traditionally human role – they tend to behave as if they were interacting with another person. As a result, common social rules and norms are applied automatically and unconsciously (Nass & Moon, 2000; Reeves & Nass, 1996). This phenomenon is described by the media equation theory, which stands for "media equals real life" (Reeves & Nass, 1996; p. 5). More specifically, the social cues described above draw people's attention away from contradicting information – such as that the artificial interaction partner is neither alive, nor able to have own feelings, thoughts, or intentions – and instead trigger social scripts, expectations, and labels (Nass & Moon, 2000). The founders of the media equation theory explain these reactions by the pronounced social nature of humankind, which leads people to rather treat something falsely as human than treating something falsely as non-human (Reeves & Nass, 1996). So far humans lived in a world where only other humans exhibited rich social behavior, which caused our brain to learn to automatically react to social cues in a certain way (Nass & Moon, 2000).

Empirical evidence for this theory shows that people react politely (Nass et al., 1999), apply gender-based stereotypes (Nass et al., 1997), and respond to flattery (Fogg & Nass, 1997b) when interacting with a computer just as if they were interacting with another person. Extending these first findings regarding interactive computers, the media equation phenomenon was also examined with smartphones (Carolus et al., 2019), virtual agents (Hoffmann et al., 2009; Kang et al., 2008), and robots (Eyssel & Hegel, 2012; Horstmann et al., 2018; K. M. Lee et al., 2006; Mumm & Mutlu, 2011; Schermerhorn et al., 2008; Woods et al., 2005). For instance, Hoffmann et al. (2009) found that people also show polite behavior when giving direct feedback to a virtual agent regarding its performance and Eyssel and Hegel (2012) found that people apply gender stereotypes to robots. In another study by K. M. Lee et al. (2006), participants attributed personality to a robot and preferred interacting with a robot which was perceived to resemble their own personality.

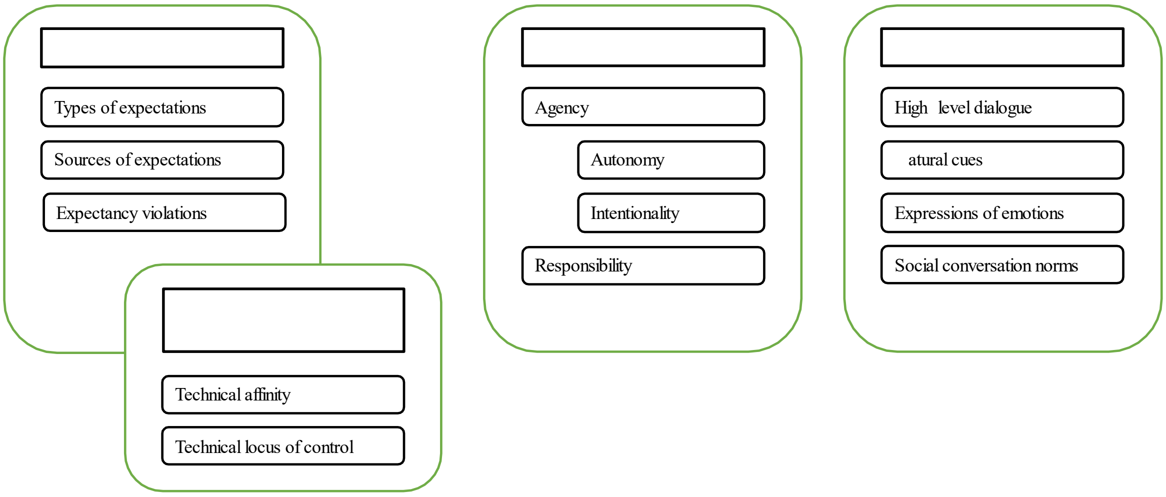
However, there is also empirical evidence for differences regarding the way people react to human and artificial entities which suggests that artificial entities are perceived differently in some regards. Results of a meta-analysis by Fox et al. (2015) indicate that person-controlled avatars are more influential than computer-controlled agents in social situations. For instance, people perform worse on a novel task in the presence of other people's avatars than alone, which was not the case with computer agents (Blascovich et al., 2002). In a virtual Blackjack game, how high people placed their bets differed when their opponent was supposedly a human or a computer program (Blascovich et al., 2002). In a negotiation situation, computer programs are evaluated more cooperative and punished less than human opponents (Gratch et al., 2016) and with a computer-controlled clinical interviewer people report lower fear of self-disclosure, display emotions more intensely, and appear more willing to disclose compared to when it was controlled by a human (Lucas et al., 2014).

Summing up, research provided much evidence for astonishing similarities regarding the perception of artificial and human entities. Although there appear to be some differences as well, theories from social psychology should work as valuable framework when examining interactions with artificial entities, as suggested by Krämer et al. (2012). Therefore, expectations alongside with technological expertise, attribution of causality and intentionality, as well as interactional behavior are considered as influencing factors for the perception of artificial entities within this thesis.

2.4. Influencing Factors for the Perception of Artificial Entities

In the following it is outlined how factors which were found to influence perceptions of other humans may also play a substantial role in interactions with artificial entities. First of all, expectations are able to shape perceptions of the reality (Burgoon, 1993; Burgoon & Le Poire, 1993), i.e., the real world may be perceived differently depending on the prevalence and salience of respective expectations. From today's perspective, personal contacts with artificial entities are still scarce (Banks, 2020). However, the sector of artificial interaction partners has a high expected growth rate (Visti, 2018), which against the background of lacking personal experiences highlights the necessity to examine what expectations people have regarding artificial interaction partners and how these are formed and influenced. Second, the user's technological expertise is considered as factor influencing the formation of expectations regarding artificial entities and consequently how they are perceived. Furthermore, since artificial entities are usually situated in real-world environments which they influence with their actions, it is crucial to consider attribution processes regarding the causality and intentionality of their behavior (Kelley, 1973; Malle et al., 2001). Attribution of agency is considered as prerequisite for these attributions and related to the perceived autonomy, intentionality, and responsibility of artificial entities (Sullins, 2006). These aspects should ultimately have a great impact on how artificial entities are perceived during and after an interaction (Kelley, 1973; Sullins, 2006). Apart from that, behavior in general was found to have a striking effect on people's evaluations of artificial entities (Rickenberg & Reeves, 2000) and is therefore considered as further influencing factor for the perception of artificial entities. An overview of all the determinants and underlying concepts – as presented in the subsequent theory sections – is shown in Figure 1.

Figure 1. Theoretical overview of the influencing factors and underlying concepts for the perception of artificial entities



2.4.1. Expectations Influencing the Perception of Artificial Entities

In general, expectations determine how an interaction partner (Burgoon & Hale, 1988) as well as technology (Davis, 1989; Sundar et al., 2016) is perceived. Since an artificial entity is a technology functioning as interaction partner, it represents both of these concepts and expectations should certainly affect its perception. To examine this is crucial because how the artificial entity is perceived will ultimately affect whether this technology will be accepted and used in the future (Davis, 1989, 1993). Furthermore, we will learn more about how expectations form human perceptions in general and when interacting with artificial entities in particular.

2.4.1.1. Types of Expectations

As described before, expectations emerge from people's need to reduce uncertainty by gathering available information (Berger & Calabrese, 1975) and forming first impressions (Asch, 1946). For the purpose of explaining as well as developing reliable predictions about others' behaviors, social categories are used (Andersen & Klatzky, 1987). Since previous research has repeatedly shown that similar processes take place in interactions with artificial entities as with humans (Eyssel & Hegel, 2012; Hoffmann et al., 2009; Horstmann et al., 2018; Reeves & Nass, 1996), there is reason to presume that people also strive to reduce uncertainty, gather information, receive a first impression, and assign a (social) category when interacting with artificial entities. However, the last step might be particularly challenging since artificial entities are clearly of technological nature but, in contrast to other technological devices, are often perceived as being alive (Kahn et al., 2011; Rosenthal-von der Pütten & Krämer, 2015; Severson & Carlson, 2010). Therefore, since no clear category can be assigned and personal experiences are rare, forming expectations about artificial entities should create difficulties. In line with these assumptions, a study showed that people expect higher levels of uncertainty when anticipating an interaction with a robot compared to a human (Spence et al., 2014). Nevertheless, due to people's need to reduce uncertainty and increase predictability, expectations will still be formed using sources of information that are available such as contextual information as well as mass media reports and science fiction formats.

In line with that, previous research showed that the appearance and behavior of an artificial entity is generally expected to be appropriate for the task context (e.g., playful for entertainment tasks, authoritative for serious tasks; Goetz et al., 2003). Here, expectations appear to be formed based on the context. Regarding characteristics of artificial entities, people predominantly name performance-oriented traits, such as efficient, reliable, and precise, in contrast to socially oriented traits, such as feeling, compassionate, and social (Ezer et al., 2009). A large-scale survey extends these findings by showing that robots are expected to be rather precise, reliable, rational, and perfectionist and rather not alive, human, and able to feel (Arras & Cerqui, 2005). Thus, artificial entities are generally not expected to possess too many human-like personality or character traits (Dautenhahn et al., 2005). With regard to their areas of application, social robots are expected to do simple, impersonal, noncreative, and repetitive household tasks such as cleaning or moving things, rather than tasks involving social aspects such as child or pet care (Khan, 1998; Ray et al., 2008).

In general, there mainly appear to be two prominent prospects for artificial entities (Bruckenberger et al., 2013; Scopelliti et al., 2005). On the one hand, there is this negative view on artificial entities becoming competitors to humans (Scopelliti et al., 2005). People are worried about too much autonomy and loss of control, which is often accompanied by a fear that humans will be either replaced or dominated by artificial entities (known as "Frankenstein Syndrome"; Asimov, 1947; Ray et al., 2008; Rosenthal-von der Pütten & Krämer, 2015; Weiss et al., 2011). Mass media and particularly science fiction has a great influence on people's image of social robots and promotes the idea of robots developing their own agenda and revolting against humans (Khan, 1998). An artificial entity which is expected to learn from humans to become better and more efficient than them in order to possibly replace them one day, for instance, at work, should be perceived as threatening or at least undesirable. On the other hand, there is this positive image of social robots functioning as assistants (Scopelliti et al., 2005), in domestic, public, or work environments (Khan, 1998; Ray et al., 2008). The idea of having an electronic help, which makes life easier by carrying out tasks that are unpleasant or strenuous, is very appealing to most persons (Oestreicher & Eklundh, 2006; Ray et al., 2008). Consequently, an artificial entity which is expected to assist humans with various tasks should be perceived very positively and desirably.

Besides the studies presented above, not much research explicitly examined expectations regarding artificial entities. A first, relevant step is to examine how people form their expectations regarding artificial entities in order to receive a better understanding about the content of their expectations. For this purpose, it is crucial to explore where and how people have been in contact with artificial entities, particularly whether this were personal encounters or receptions of media portrayals.

2.4.1.2. Sources of Expectations

One source for expectations regarding artificial entities are personal encounters, for instance, at people's workplaces, at exhibitions, or as part of an experimental study. These kinds of encounters, however, are not very common yet and many people have not personally interacted with a social robot or virtual agent so far (Banks, 2020; Bruckenberger et al., 2013). Consequently, other sources than personal experiences need to be used to reduce uncertainty and develop expectations, whereby mass media bear a great potential (Banks, 2020).

Mass media are likely a popular source since they are widely accessible and reach a great audience (Nisbet et al., 2002). As previous research by Gerbner and Gross (1976) showed, media content is generally able to influence how people perceive the real world. Their studies revealed, for instance, that people who are heavy TV viewers overestimate the proportion of people employed in law enforcement and are generally more careful in trusting people compared to light TV viewers. Particularly, media were found to affect perceptions of technologies (Nisbet et al., 2002) and to be important means by which people can seek information about a technology with which they had no personal contact yet (Kriz et al., 2010; Sundar et al., 2016). With regard to artificial entities, mass media portrayals should be divided into reports/documentaries about real existing technologies and science fiction movies or series presenting fictional robot characters.

Although reports and documentaries may depict actual existing artificial entities, they still often fail to portray uncertainties, unsolved technological problems, and setbacks adequately, while enhancing successes and progresses of a technology (Weiss et al., 2011). In other words, the advanced site of artificial entities is presented to the general public, while shortcomings and deficiencies are neglected (Weiss et al., 2011). Thus, reports about existing artificial entities appear to be biased in a way that leads to an overestimation of their skills.

In case of science fiction formats this bias is even more pronounced (Bruckenberger et al., 2013; Sandoval et al., 2014). Artificial entities often play a central role in science fiction (e.g., Asimov, 1947, 1950; Clarke, 1968; Dick, 1968), which is believed to have a remarkable effect on people's expectations (Khan, 1998; Weiss et al., 2011). What people expect, in turn, will affect how they perceive these technologies at present. For instance, the well-known science fiction author Isaac Asimov (1950) presented the Three Laws of Robotics: minimizing harm to humans, self-preservation, and obeying orders. These laws address what people often fear the most with regard to robots in the future (McCauley, 2007). Particularly, people with no personal experiences with artificial entities often use science fiction as reference (Kriz et al., 2010; Ray et al., 2008; Scopelliti et al., 2005) and emphasize how much their knowledge of and attitude towards artificial entities are affected by science fiction portrayals (Bruckenberger et al., 2013; Weiss et al., 2011). In science fiction, robots are usually equipped with advanced humanlike cognitive skills, which were found to positively correlate with people's expectations regarding the cognitive capabilities of real-world robots (Kriz et al., 2010). Despite the fact that the physical and cognitive abilities of actually existing artificial entities are far behind those portrayed in science fiction (Murphy, 2018), people's expectations, particularly regarding skills and abilities, but also regarding the general scope of application for artificial entities, appear to be greatly affected by their portrayals in mass media. As a result, heightened expectations emerge which may be violated as soon as an interaction with an actual artificial entity occurs which does not possess the skills portrayed in mass media.

2.4.1.3. Expectancy Violations

Expectations describe possible, feasible, appropriate, and typical behaviors for a certain composition of factors related to the situation, the interaction partner, and their relationship (Burgoon & Hale, 1988). People always try to make reliable predictions about their interaction partners. However, the partner's behavior still may deviate from expectations, which results in so-called expectancy violations (Burgoon & Hale, 1988). As outlined by the expectancy violations theory (Burgoon & Hale, 1988; Burgoon & Jones, 1976), behavior can deviate by being worse than expected, resulting in a negative expectancy violation, or by being better than expected, resulting in a positive expectancy violation. Negative violations are usually followed by detrimental consequences, while positive violations appear to be accompanied by favorable ones, which exceed positive reactions to confirmed expectations (Burgoon & Hale, 1988).

In its original form, the expectancy violation theory was developed to explain why changes in conversational distance during a social interaction were found to have different outcomes depending on the interaction partners and expectations related to them (Burgoon & Hale, 1988). Since then, the model has been revised and extended to be applied to other behaviors, situations, and communication outcomes (Burgoon & Hale, 1988). Recently, the theory was even applied to interactions with virtual agents, where results showed that a positive expectancy violation can result in higher task attractiveness ratings compared to a negative violation or an expectancy confirmation (Burgoon et al., 2016).

Since previous research generally encourages the appliance of social psychological theories to interactions with artificial entities (Reeves & Nass, 1996), the expectancy violations theory may provide a useful lens for explaining and predicting human perception and behavior when interacting with artificial entities (Burgoon et al., 2016). As mentioned before, people expect sophisticated abilities, efficiency, and precision from artificial entities (Arras & Cerqui, 2005; Ezer et al., 2009). Due to a lack of personal experiences, these expectations are predominantly based on representations in mass media (Banks, 2020), which likely results in rather high expectations regarding the skills of artificial entities and their future role in people's daily lives (Bruckenberger et al., 2013; Kriz et al., 2010; Weiss et al., 2011). Consequently, the occurrence of expectancy violations is highly probable.

Previous research already addressed the problem of unmet or violated expectations in human-robot interaction, which are described as gaps between people's expectations and their actual perception of a robot's function (adaption gap: Komatsu & Yamada, 2011; expectations gap: Kwon et al., 2016). Similar to the principles of the expectancy violations theory, Komatsu and Yamada (2011) argue that the difference between a user's preexisting expectations regarding the functions of a robot and how the functions are then perceived during an interaction will significantly affect the user's behavior toward and acceptance of the robot. Moreover, the authors also distinguish between positive and negative adaption gaps, depending on whether a perceived function exceeds or falls behind an expected function. It is further assumed that a negative gap leads to disappointment and a reduction of trust in the robot, while a positive gap has opposite effects. According to Kwon et al. (2016), people tend to generalize social capabilities for humanoid robots, which results in unrealistically high expectations of social robots and may consequently lead to disappointment, mistrust, and rejection.

2.4.2. Technological Expertise Influencing the Perception of Artificial Entities

Expectations regarding artificial entities may be predominantly formed by personal experiences or mass media, however, people's general expertise regarding technology may also have an influence. Since artificial entities are of technological nature, people's general enthusiasm about new technologies and how confident they feel when handling technologies should have an impact on the expectations people develop. Ultimately, this should also have an effect on how people perceive artificial entities as technologies. Thus, in the following, technical affinity as well technical locus of control are looked at more closely.

2.4.2.1. Technical affinity

Enthusiasm towards technology is described by the term technical affinity or technophilia (Karrer et al., 2009). Technical affinity represents how much a person is attracted or positively inclined to particularly new technologies. People with a high technical affinity usually enjoy to explore different types of applications and eagerly try new and even complex technologies (Morahan-Martin & Schumacher, 2007). Technical affinity should impact how artificial entities are perceived since they pose an exciting, innovative as well as complex technology. Due to their overwhelming enthusiasm, people with a high technical affinity potentially perceive artificial entities more positively while tending to overestimate how capable they are at the present and will be in the future.

2.4.2.2. Technical locus of control

How competent a person feels when using technologies is described by their technical locus of control (Beier, 1999). When a person feels capable of reaching a certain goal by using technology, this person's perceived locus of control regarding the usage of technology is high. When a person however feels helpless and overwhelmed when handling technological devices, their perceived locus of control in this regard is low (Gaul et al., 2010). Whether a person feels confident or helpless regarding the usage of technology is assumed to be affected by their actual technological competence and knowledge (Beier, 1999). As a consequence, people with a high locus of control when using technologies may have a more realistic view on the capabilities of artificial entities. This should on the one hand affect what they expect with regard to artificial entities and on the other hand how they will perceive those technologies as interaction partners.

2.4.3. Attribution Processes Influencing the Perception of Artificial Entities

As outlined before, attribution processes take place in interpersonal settings to determine whether the cause for an action can be found within the acting person and/or their environment (Kelley, 1973). Attributing behavior to a person's disposition is referred to as internal cause and attributing behavior to situational or contextual factors as external cause (Kelley, 1973). Furthermore, the attribution of intentionality needs to be taken into account as well, since behavior that is considered intentional is also evaluated based on the motives that are perceived to have guided that behavior (Malle et al., 2001; Reeder et al., 2002). Considering these findings from interactions between humans, the question arises whether the behavior of artificial entities is rather attributed to internal or external causes as well as whether and what kind of motives are assumed to form the basis for their behavior. Looking at it from a logical point of view, the behavior of artificial entities is caused by external factors since they are technology-based, restricted to their programming, and not able to have a personal disposition, at least not shaped by own intentions, emotions, needs, and attitudes (Bigman & Gray, 2018; Malle & Knobe, 1997). In short, the behavior of artificial entities should be perceived to be caused externally by its programming and not be driven by own intentions. However, perceptions of artificial entities do not necessary follow logical deliberation (Reeves & Nass, 1996). Therefore, research looking at the perception of factors which determine whether an artificial entity's behavior is attributed to the entity and its intentions or surrounding circumstances is of great relevance.

2.4.3.1. Agency

A prerequisite for being perceived accountable is the capability to perform self-directed actions called agency (Himma, 2009), a term that is extensively discussed in human-machine interaction. There are various definitions for agency focusing on different aspects, contexts, and relations (Banks, 2019; Emirbayer & Mische, 1998). For this work, agency is understood as the capacity to act independently and free of choice (Barker, 2005), i.e., as a form of self-motivated governance (Banks, 2019). In addition to free will (Allen et al., 2006), self-motivation (Emirbayer & Mische, 1998), and animacy (Brown & Walker, 2008), consciousness is named as precondition. All of these prerequisites would make agency clearly ascribable to natural agents such as humans in contrast to artificial entities such as robots and virtual agents (Himma, 2009). In line with that, results of a meta-analysis conducted by Fox et al. (2015) show that more agency is reported with human-controlled than computer-controlled entities.

However, people are often observed to react to artificial entities on an unconscious level as if they were alive (Reeves & Nass, 1996), which is why it might be more pivotal to focus on whether people perceive an artificial entity to possess agency while neglecting whether it is able to possess it by definition. Particularly since artificial entities are becoming progressively interactive and take over human roles, perceived moral agency needs to be reflected upon as well (Banks, 2019; Sullins, 2006). Moral agency combines the terms morality and agency (Banks, 2019) and thus results in the question whether an artificial agent is perceived to be able to act (im-)moral. In the following necessary prerequisites for the perception of moral agency are illuminated which are, according to Sullins (2006), autonomy, intentionality, and responsibility. These factors should play an important role for the internal or external attribution of an artificial entity's behavior, as well as for the attribution of motives, which should affect how this behavior and consequently the entity are perceived.

2.4.3.2. Autonomy

Autonomy is related to an internal locus of causality, which means that the variables which cause an action are to be found within the acting entity (de Charms, 1983; Deci & Ryan, 1987). In other words, the perceived origin or force of the artificial entity's actions is shifted from the human controller to the artificial entity itself (Gunkel, 2014). In general, autonomy results from the independence of control by any other agent, such as programmers, manufacturers, or owners (Banks, 2019; Sullins, 2006).

Full independence is hard to achieve, not only for artificial entities but also for living beings since there are always factors which influence one's actions (Sullins, 2006). For artificial entities these factors may be forces exerted by the technology, the system, developers, or social interaction partners (Sullins, 2006). However, as argued before with agency, the focus should be put on the effects of perceived autonomy rather than actual autonomy. Whether an artificial entity is perceived as acting autonomously on its own or being controlled by an external force, should impact how its actions are perceived. For instance, during an experiment by T. Kim and Hinds (2006), people attributed more blame to a robot whose behavior suggested high autonomy than to themselves or other participants compared to when the robot's behavior indicated low autonomy.

Looking at the term 'autonomy' from an engineering point of view, it can be defined as being under no direct control of another agent or user, while not excluding being influenced by another agent or user (Sullins, 2006). In this vein, Bartneck and Forlizzi (2004) explain that the autonomy of a social robot describes its technological capability to act without direct input from a human. Consequently, an avatar which is defined as "virtual representation of a human" (von der Pütten et al., 2010, p. 1641) and fully controlled by a person in real-time, should not be perceived as acting autonomously. Besides direct input by a human, a program may also exert direct control over or just influence the behavior of an artificial entity, which could determine whether it is perceived as acting autonomously or not. For instance, the interactional behavior of an artificial entity could be fix-programmed in advance (if interaction partner says X, then answer Y). During the interaction, the entity's behavior would be fully controlled by this program, which would leave no room for autonomous actions. However, an artificial entity that was trained to generate an appropriate answer by itself, based on what an interaction partner says, is influenced but not controlled by its programming. In a sense, this artificial entity would act with more freedom and may consequently be perceived as more autonomous.

Summing up, an important component of agency is autonomy, i.e., the ability to act independently of forces exerted by another agent, which could be a person or computer program. When another entity is controlling the behavior of the artificial entity, the agency may be shifted away from the artificial entity to the person or program which exerts the control. This should ultimately affect where the behavior is attributed to.

2.4.3.3. Intentionality

Another important aspect for agency is intentionality, i.e., whether the agent is following its own, someone else's or no one's intentions while acting. According to Breazeal and Scassellati (1999), artificial entities need to convey intentionality in order to interact socially with humans. Intentionality is described as the ability to act with volition resulting in deliberate actions (Banks, 2019). As described by Banks (2019) "an intentional entity means to do what it does, independent of rote programming or conditioning" (p. 364). Intentionality is usually the basis for intrinsically or extrinsically motivated behavior, i.e., behavior to attain a certain outcome which is either the behavior itself or separate from the behavior (Deci & Ryan, 1987). The distinguishing feature of intentions is that they comprise an action content while desires, for instance, may have any content such as one's own as well as someone else's behavior or a certain event or outcome (Malle et al., 2001). The degree to which intentions determine an agent's actions is affected by the amount of control an agent has over their actions (Fishbein & Ajzen, 2010). A lack of internal resources such as knowledge, skills, and abilities as well as external impediments may hinder agents to carry out their intentions (Fishbein & Ajzen, 2010).

After defining what intentions are and how they determine an agent's actions, the question remains how they emerge. In general, intentions are perceived to be the immediate precursors of a person's behavior (Lillard, 1998).When people are asked to define intentionality, mainly four components are mentioned: desire (for a certain outcome), belief or knowledge (about the action to attain that desired outcome), intention (to perform the action), and awareness (of performing the action; Malle & Knobe, 1997). Thus, intentions appear to be generally born from desires, which may stem from certain intrinsic needs and motivations (Banks, 2019).

From a logical point of view, artificial entities are not able to have intrinsic needs, motivations, or desires and thus are not capable to possess own intentions. Instead, their actions are either determined by a person or a program. Nevertheless, based on the insights of the media equation theory (Reeves & Nass, 1996), people may perceive the actions of artificial entities to stem from own intentions based on personal desires, needs, and motivations. For instance, a study by Krämer et al. (2017) showed that a human and a computer interaction partner are both evaluated negatively after providing participants with negative feedback. This indicates that negative feedback is perceived to be intentional and attributed to the feedback giver regardless whether it is a person or a computer. Studies further show that even abstract geometrical shapes that move around on a computer screen can be perceived as acting with intentions, particularly if they seem to deliberately interact with each other and their environment (Blythe et al., 1999; Heider & Simmel, 1944; Scholl & Tremoulet, 2000). In line with that, Bartneck and Forlizzi (2004) pointed out that an autonomously acting social robot could be perceived as acting on its own intentions, although it does not have volition like humans do. When the behavior itself suggests that it is planned and executed with a goal, people are quick to assume and attribute intentionality. Malle et al. (2001) may deliver an explanation for that by pointing out that the attribution of intention allows people to detect structure in the behavior of others, which could reduce uncertainty. Thus, in order to be able to explain and predict an artificial entity's behavior, people may be inclined to ascribe intentionality (Berger & Calabrese, 1975).

However, there are also circumstances under which computers are perceived to be acting with less intentionality than people. This was shown to be the case with the ultimatum game, where a fixed amount of money needs to be divided between two players by accepting or rejecting a splitting proposal. In case of a rejection, none of the parties receive any money (Blount, 1995). When a person proposes how to split the money, people appear to expect fair offers. An unfair distribution is perceived to be intentional and usually rejected as a form of punishment accompanied by anger and disgust as answer to the perceived violation of social norms (Blount, 1995; Chapman et al., 2009). When a computer however is suggesting how to split the money, more uneven distributions, i.e., lower amounts of money, are accepted (Blount, 1995). This suggests that the computer is perceived to be less intentional when proposing those unfair splits which appears to result in less judgements and a lower perceived need to punish the computer by not accepting the offer. Nevertheless, people still do not accept all offers which they should if they would perceive the computer as acting with no intention at all. Thus, an attribution of intentionality to an artificial entity still appears to take place, although to a lesser extent than with a human interaction partner.

In general, the degree to which autonomy and intentionality are perceived to be present when interacting with an artificial entity should affect how much responsibility is ascribed to the artificial entity for the way it behaves. As a study showed, people reported to feel less responsible when collaborating with a humanlike than a machine-like robot (Hinds et al., 2004). An explanation could be that the human-like robot was perceived as more autonomous as well as intentional and thus more capable of carrying responsibility than a machine-like robot (Hinds et al., 2004). Perceived autonomy and intentionality appear to be determinants for the perception of agency. Together with perceived responsibility, agency should decide over the degree an artificial entity's behavior is attributed to the entity itself or external factors, which should then be decisive for the overall perception of the artificial entity.

2.4.3.4. Responsibility

As outlined before, responsibility results from situational and role-related factors such as a teacher being in charge in a classroom or a captain being in charge on a boat. Artificial entities are able to take over those kind of roles (Nass & Moon, 2000) and accordingly should be perceived responsible when doing so (Sullins, 2006). For instance, a social robot taking over the role of a care giver for elderlies should be perceived responsible for executing important tasks reliably such as reminding patients to take their medicine. If the robot fails to remind its patients to take their medicine, the robot will be blamed for this failure, but only if the robot was perceived to be acting autonomously as well as intentionally and was perceived to be responsible. Thus, perceived autonomy and intentionality appear to be prerequisites for perceiving this entity to act with agency. Behaviors and their consequences are then attributed to an artificial entity if it is perceived to possess agency and to be responsible. After discussing all these prerequisites, let us take a look at the circumstances influencing whether and to which degree behaviors and their consequences are be attributed to an artificial entity.

2.4.3.5. Attribution Processes in Interactions with Artificial Entities

Since the technologies behind artificial entities are advancing to become increasingly reliable, more and more tasks are delegated to them to be performed autonomously (Gogoll & Uhl, 2018; Hancock, 2017). This raises the question how actions performed by these increasingly autonomous artificial entities are perceived, particularly whether they are attributed to the entity or other external forces such as programmers, developers, or manufacturers. Artificial entities may never act fully autonomously and intentionally or be able to carry full responsibility, however, they might be perceived to meet those criteria. Accordingly, the focus should be laid on the perception of agency and responsibility in order to evaluate whether behaviors and subsequent consequences are attributed to the artificial entity displaying them.

In case of an avatar or a robot that is controlled in real-time and immediately transfers the person's intentions, the behavior should be clearly attributed to this person. The avatar or robot would just be a computer-graphical or technological representation of that person and one would interact with the person through the representation. This would rather fall into the category of computer-mediated communication than human-computer interaction (Fox et al., 2015; Morkes et al., 1999) and is thus not further reflected upon here.

A virtual agent or a robot that is acting based on a program, though, is also not acting completely autonomously or intentionally since it was programmed by a person or a group of persons at some point to act this way. However, research showed that programmers do not play a particular role in people's perception of and reactions to artificial entities (Nass & Moon, 2000; Sundar & Nass, 2000). Thus, when interacting with such an artificial entity, people neglect that it was programmed to act this way and rather attribute its behavior to be caused by itself. As discussed before, people tend to react to artificial entities not in a logical but in a social way (Reeves & Nass, 1996). Thus, behaviors may generally be attributed to factors located inside of the artificial entity, despite the fact that artificial entities have no free will over their actions the way humans do. Although people may state to be aware that artificial entities are not able to possess human-like traits such as own intentions, emotions, and attitudes when asked directly, they still may react to these entities during an interaction as if they would possess those traits (Nass & Moon, 2000).

A new hybrid form of artificial interaction partner called agent-representative poses a special case since it mixes and blurs computer- and person-controlled agency. Agent representatives are artificial entities, which are supposed to represent a specific person and their interests in a certain situation (de Melo et al., 2016). These agents are acting autonomously in the situation but receive specific instructions on how to act by the person they are supposed to represent, so that they may act in accordance with the person's intentions, motivations, beliefs, and attitudes. Examples for agent representatives which are already in use today are automated phone call systems such as Google Duplex (Leviathan & Matias, 2018) or bidding systems as known from the Internet auction platform eBay as well as automated negotiators and self-driving cars (de Melo et al., 2018). Since agent representatives only act upon a person's intentions, the instructing person should be held responsible for the behavior of the agent representative. However, to whom the cause of the agent representative's behavior needs to be attributed to might not be clear and blame for the negative behavior may be deflected (Bivins, 2006; Royzman & Baron, 2002) or even attributed to the agent representative itself.

In general, the life-like behavior of artificial entities combined with a lack of transparency regarding what causes their behavior may lead people to perceive artificial entities as acting upon own intentions and following an own agenda (Bartneck et al., 2009; Bartneck, Kanda, et al., 2007). The perception of animacy or life-likeness in artificial entities is enhanced, for instance, by their physical behavior (Bartneck et al., 2009; Bartneck, Kanda, et al., 2007; Bartneck, van der Hoek, et al., 2007), their communication and social skills (Dautenhahn, 1999; Ishiguro et al., 2001), their human-like appearance (Riek et al., 2009), as well as by assigning personal names, stories and experience (Darling et al., 2015). However, in general not many cues are necessary to make humans react to artificial entities as if they were alive.

As mentioned before, people generally tend to overestimate internal factors regarding a person's disposition, while underestimating external, situational factors – a phenomenon known as fundamental attribution error (Jones & Nisbett, 1972; Nisbett & Ross, 1983; Ross, 1977). Calling back to mind the study by Jones and Harris (1967), where people perceived a position that was assigned to people via coin toss to reflect their actual attitude, it needs to be considered that similar processes may take place in human-machine interaction as well. In this vein, external constraints of an artificial entity such as a fix programming or strict instructions by a person may be neglected. This would lead to perceiving the artificial entity itself as fully responsible, which would have a great impact on the perception of the artificial entity.

Summing up, attribution processes are linked to the extent to which an artificial entity is perceived to possess agency – based on perceptions of autonomy and intentionality – as well as to be responsible – based on situational and role-related aspects (Sullins, 2006). Examining against this background whether an artificial entity's behaviors and subsequent consequences are attributed internally to the artificial entity itself or externally in consideration of constraining factors such as a fix programming or instructions should have a great impact on the perception of these behaviors (Kelley, 1973; Kelley & Michela, 1980). How the behavior of an artificial entity is perceived should consequently determine the overall perception of the artificial entity. During an interaction, the behavior of an artificial entity may pose the most influential factor for people’s perception of the artificial entity (Rickenberg & Reeves, 2000), which is reflected in the subsequent section.

2.4.4. Behavior Influencing the Perception of Artificial Entities

According to Reeves et al. (2020), artificial entities are “programmed to behave in variety of social or helpful ways, enough so that the people who use them could interact via familiar and pleasurable and even emotional connections” (p. 1). Animated artificial entities are considered to make social interactions more pleasing, productive, and easy (Rickenberg & Reeves, 2000). This was found to bear several benefits, for instance, when using artificial entities as pedagogical agents (Lester et al., 2000; Paiva & Machado, 1998) or as support for elderlies (Heerink et al., 2009; Looije et al., 2010; Tanaka et al., 2017). Behavior is a wide-ranging term which includes but is not limited to body movements, gaze directions, facial expressions, voice and intonation, postures, gestures, and other verbal as well as non-verbal expressions. An overview over the various behaviors of artificial entities is given in the following.

2.4.4.1. rtif icial nti ties’ Behavior Variations

In general, an artificial entity’s behavior can be designed to show high or low amounts of expressivity and sociability (Heerink et al., 2009). For this purpose several behavioral aspects such as gaze, response to errors, expressions of emotions such as smiling and cheerfulness, recall of personal details such as the user’s name, nodding, blinking, turn-taking, and politeness can be means to manipulate people’s perception of an artificial entity (Breazeal, 2003; Cassell & Thórisson, 1999; Heerink et al., 2009; Looije et al., 2010). Against this background, Looije et al. (2010) distinguish between four behavior types, which are outlined in detail below: high-level dialogues, showing natural cues, expressing emotions, and social conversation norms.

With regard to dialogue, an artificial entity may use strong language containing many confident statements or weak language including questions and suggestions (Isbister & Nass, 2000). An artificial entity can be rude, selfish, and bossy as well as kind, pleasant, and empathetic (Heerink et al., 2009; Mumm & Mutlu, 2011). According to Dautenhahn (2007), the reliability and robustness of interaction skills may vary, which should result in different levels of dialogue quality. Concerning voice and intonation, the volume, speech rate, frequency as well as frequency range may be adapted (S. Lee et al., 2005; Nass & Lee, 2001). Moreover, the artificial entity can be equipped with a rather machine-like or rather human-like voice (Baylor et al., 2003), which further can be female or male (Eyssel et al., 2012).

Concerning natural cues, the gaze direction of an artificial entity may be directed to or averted away from a human interaction partner and/or a shared target (Bruce et al., 2002; Cassell & Thórisson, 1999; Heerink et al., 2009; Kozima et al., 2003; Mumm & Mutlu, 2011). With regard to body movements, an artificial entity may display proxemic behavior, i.e., approach or keep distance to a person (Isbister & Nass, 2000; Satake et al., 2009). Furthermore, different forms of gesticulation can be implemented such as wide and free or narrow and restricted (Isbister & Nass, 2000).

Emotions can potentially be expressed by artificial entities via body movements such as gestures and postures (Embgen et al., 2012; Häring et al., 2011; Pelachaud, 2009b; Tan & Nareyek, 2009), sounds (Häring et al., 2011), eye colors (e.g., red for anger or negative valence; Häring et al., 2011; Rosenthal-von der Pütten et al., 2018), verbal expressions (Horstmann et al., 2018), as well as facial expressions (Breazeal, 2002; Bruce et al., 2002; Kühnlenz et al., 2010; Pelachaud, 2009a; Tan & Nareyek, 2009). For instance, Bittermann et al. (2007) used a robotic head which is able to express six basic emotions and Ochs et al. (2010) examined the effects of different types of smiles by a virtual agent.

Regarding social conversation norms, turn-taking is an important aspect which many researchers recognized and chose as research object (e.g., Breazeal, 2003; Cassell & Thórisson, 1999; Looije et al., 2010; Skantze et al., 2014). Other social conversation norms may be politeness (Carolus et al., 2019; Nass et al., 1999), synchrony and reciprocity (Lorenz et al., 2016), as well as nonverbal behaviors to give feedback during an interaction, such as nods and glances (Cassell & Thórisson, 1999).

All of the above are behavior examples to show how complex and multi-faceted behaviors of artificial entities can be. Consequently, effects of different behavior variations on their human interaction partners are similarly complex. Many researchers made attempts to examine those effects, however, extensive research is still needed. Since it is not even fully researched yet how people respond to behaviors by other people (a goal which may never be reached due to its complexity), it is only logical that there are also still big gaps with regard to interactions with artificial entities. In the following, some of the findings regarding the effects of behavior on the perception of artificial entities are presented.

2.4.4.2. Effects of rtificial nt ities’ Behaviors

Along with the many possibilities to display different behaviors, there is an extensive body of research examining how people perceive, evaluate, and react to these behavioral variations when interacting with artificial entities. In general, an interaction with artificial entities accompanied by behaviors such as gaze, manual beat gesture, and head movements is preferred over an interaction missing those behaviors (Cassell & Thórisson, 1999). There are several behaviors which can be perceived as socially interactive such as avoidance of collisions and interruptions via anticipatory measures, showing interest, moving the head or eyes meaningfully towards a person or a shared object of interest, turn-taking, and active participation (Dautenhahn, 2007). However, certain behaviors which indicate that the artificial entity is taking no interest in what the human interaction partner is doing or saying can be perceived as socially ignorant (Dautenhahn, 2007). Particularly, this could be collisions with people, interruptions, being passive during a group task, and not moving the head or eyes to the human or a shared object, all of which are found to be disliked and cause discomfort (Dautenhahn, 2007). Altogether, behavior needs to be considered as having a great impact for interactions with artificial entities as further outlined below.

First of all, the media equation theory itself postulates that the behavior of an artificial entity is central for eliciting social responses (Reeves & Nass, 1996). As key factors, interactivity, natural language, and the filling of a traditionally human role are named (Nass & Moon, 2000; Reeves & Nass, 1996), which are all behavioral aspects. In this vein, the behavior of computers was shown to have an effect on people’s perceptions, e.g., in form of a better evaluation after flattering feedback (Fogg & Nass, 1997b), by responding politely in case a computer personally asks for feedback (Nass et al., 1999), as well as by applying gender stereotypes based on the voice of the computer (Nass et al., 1997).

Verbal as well as non-verbal behaviors were found to prompt people to perceive an artificial entity to possess certain personality characteristics (Isbister & Nass, 2000; K. M. Lee et al., 2006; Nass & Lee, 2001). A male or female voice can further influence how the gender of an artificial entity is perceived (Eyssel et al., 2012) and an agent with a human voice is perceived more human-like (Baylor et al., 2003; Eyssel et al., 2012). Altogether, the perceived similarity between the artificial entity and oneself appears to play a central role in the perception of these technologies (Eyssel et al., 2012; Horstmann et al., 2018). Directed gaze behavior of artificial entities is further perceived to convey attention, which elicits social behavior (Kozima et al., 2003) and makes it more compelling for people to interact with them (Bruce et al., 2002). It further depends on the artificial entity's previous behavior how people perceive and react to direct gaze by it. For example, with a rudely behaving robot, people react with an increase in distance and a decrease of self-disclosure to direct gaze by this robot (Mumm & Mutlu, 2011). In general, facial expressivity was found to make it more attractive to interact with an artificial entity (Bruce et al., 2002).

An artificial entity displaying social behavior such as turn-taking and expression of emotions is perceived as more empathic as well as trustworthy and elicits more positive responses than a text-based interface (Looije et al., 2010). For conversations with artificial entities, emotional feedback such as smiling is even topped by envelope feedback which comprises gaze and head movements (Cassell & Thórisson, 1999). With regard to movements, the mobility of an artificial entity affects how competent and warm it is perceived (Reeves et al., 2020). Further, different emotional states of an artificial entity were successfully displayed via body movements, for instance, joy via dance moves and fear via cowering (Häring et al., 2011).

Generally, it was shown that any type of non-verbal behavior, whether it is human-like or technology-specific, significantly increases the perceived animacy of the artificial entity as well as experienced positive affect and display of self-disclosure (Rosenthal-von der Pütten et al., 2018). Furthermore, employing several behavioral systems such as a combination of facial expressions, gestures, and body postures has a stronger effect than using only one of the systems (Tan & Nareyek, 2009). Summing up, it is of central importance to consider the strong impact of an artificial entity's behavior when examining which factors influence the perception of this entity. The behavior's influence may be positive, negative, or both depending on the content of the behavior and how it is interpreted (Rickenberg & Reeves, 2000). As Rickenberg and Reeves (2000) explain, "an animated character turns up the volume on social presence, which means that it can accentuate the effects of everything presented" (p. 55). Behavior therefore needs to be considered as important influencing factor for the perception of artificial entities, individually as well as in correspondence with expectations and attributions which often shape how behavior is perceived and interpreted.

2.5. Conclusion and Research Objectives

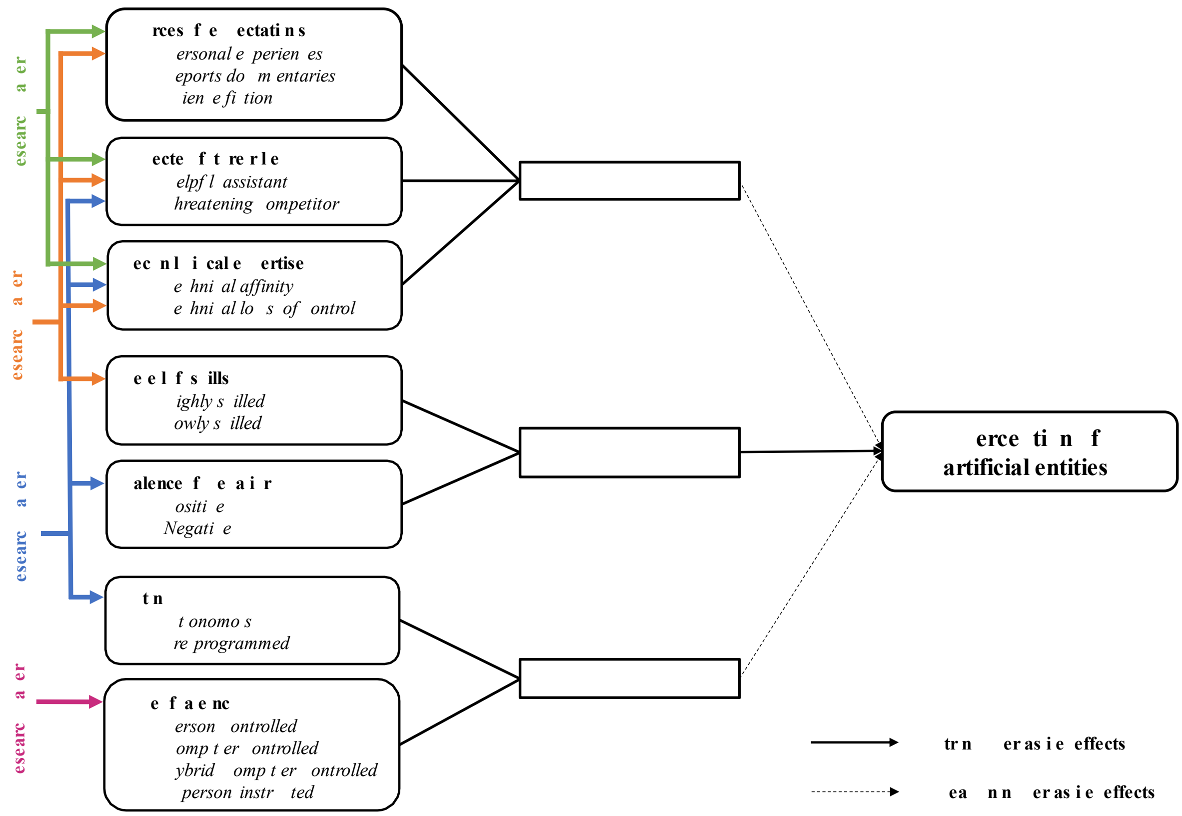
Based on the presented theoretical background, I am coming to the conclusion that expectations, attributions, as well as behavior need to be considered as influencing factors for the perception of artificial entities – individually as well as in interaction with each other. As outlined before, there is much research dealing with the effects of different behavior variations on the perception of artificial entities (see chapter 2.4.4.2). However, not much if any research considered expectations and attribution processes as further influencing factors, although they were found to have recognizable effects in human-human interactions (see chapter 2.2.1 and 2.2.1). Since many processes that were examined within interpersonal contexts were also observed to play a role in interactions with machines (Nass & Moon, 2000; Reeves & Nass, 1996), the mechanisms of expectations and attributions should be considered here as well. This is where the studies of this thesis are filling a research gap in order to understand how perceptions of artificial entities emerge and are further shaped.

Against this background, the research objectives of this thesis are to examine a) individuals' expectations regarding artificial entities alongside with their technological expertise, b) the circumstances under which an artificial entity's behavior is attributed internally or externally, c) artificial entities' behaviors in general, and d) how all of these factors influence the perception of artificial entities. Not only the individual main effects of those influencing factors are of interest, but particularly how they interact with each other in order to receive an in-depth understanding on how people form impressions of artificial entities. The studies which are contained in this cumulus and summarized in the following present different theoretical as well as methodical approaches to attend to this matter with the overarching goal to shed further light on the mechanisms influencing the human perception in general and within interactions with artificial entities in particular.

3. SUMMARY OF THE RESEARCH PAPERS CONTAINED IN THE CUMULUS

In this chapter, brief descriptions of the five studies which were conducted within the framework of the present dissertation are presented. The studies focus on different factors influencing the perception of artificial entities as portrayed in Figure 2. The pressing questions that were examined by means of these studies is how people's expectations and attributions as well as artificial entities' behaviors influence how they are perceived. The individual effects of these different factors are considered but also, with even more interest, how they interact with each other.

Figure 2. Overview of the considered influencing factors for the perception of artificial entities and the respective research papers of the cumulus



In the first two studies (research paper I), people's expectations regarding artificial entities, particularly sources of expectations such as experiences in real life and in the media as well as people's technological expertise, are examined. The third study (research paper II) employs an experimental setting to examine the influence of people's technological expertise, their expectations regarding an artificial entity's future role, and the level of the artificial entity's interaction skills on how it is perceived after an actual interaction took place. The fourth study (research paper III) investigates by means of another experimental setting the influence of an artificial entity's behavior (positive or negative feedback) and whether this behavior was portrayed as being autonomous or pre-programmed. Additionally, the influence of people's technological expertise and the artificial entity's expected future role were measured. The last study (study V, research paper IV) focuses on the role of different types of agency (computer-controlled, person-controlled or hybrid) and levels of expertise of an artificial entity on people's perception of it. Please note that all studies included in the cumulus were approved prior to the conduction by the ethics committee of the University of Duisburg-Essen.

3.1. *Research Paper 1:* Great Expectations? Relation of Previous Experiences with Social Robots in Real Life or in the Media and Expectancies Based on Qualitative and Quantitative Assessment (Horstmann & Krämer, 2019)

The research objectives of the first research paper were to explore people's expectations regarding artificial entities via qualitative interviews and to systematically examine mechanisms influencing the formation of those expectations, such as personal experiences, reception of media portrayals, and technological expertise, via a quantitative online study. During a first encounter, people experience a lot of uncertainty regarding their interaction partner and the pronounced need to reduce this uncertainty by seeking information about the other (Berger & Calabrese, 1975). Against this background, the question arises how people reduce uncertainty when interacting with artificial entities such as social robots. Since processes that take place in interpersonal settings have also been observed in interactions with artificial entities (Reeves & Nass, 1996), people will likely attempt to reduce their uncertainty by forming expectations based on available sources of information. These sources of information may be personal experiences, realistic or fictional portrayals in mass media, or general technological expertise (Bruckenberger et al., 2013). To explore these expectations and how they are formed is pivotal since they may affect people's perception (Burgoon, 1993) and consequently their acceptance of artificial interaction partners (Venkatesh & Davis, 2000). Thus, a mixed method approach was applied to achieve a more comprehensive picture in depth as well as breadth regarding people's expectations regarding social robots and how they are formed and shaped.

In a first step, thirteen semi-structured interviews were conducted to learn about people's expectations, particularly regarding applications, functions, as well as own emotions and intentions (study I). Robots are foremost expected to assist with household and work-related tasks, but not to overtake social tasks and not to have own emotions and intentions. Furthermore, fears of social robots developing an own consciousness and trying to compete with humans were mentioned frequently, which represents common science fiction scenarios and leads to the second study of this paper.

In the next step, mechanisms influencing the development of these expectancies were explored via a quantitative online study with 433 participants (study II). Focus was laid on sources of information, particularly experiences with real robots, reception of reports about real robots, and knowledge of fictional robots in science fiction movies or series. Science fiction in particular was assumed to be a central influencing factor since social robots, in contrast to science fiction formats, are currently not wide-spread and real contact is scarce (Bartneck, 2004; Bruckenberger et al., 2013; Sandoval et al., 2014). Moreover, the influence of people's technological expertise was considered as well.

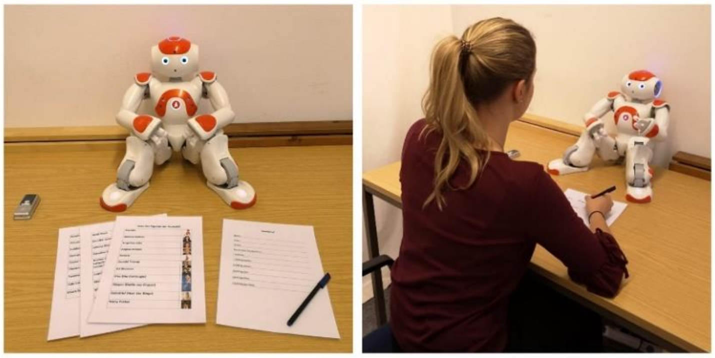
With the data attained via the online study, a structural equation model was constructed which shows that experiences with robots, particularly through mass media such as science fiction formats cause people to increasingly expect robots to have elaborated skills. To expect elaborated skills, in turn, leads people to generally expect social robots more strongly to become part of the society as well as their personal lives. Additionally, the more negatively perceived fictional robots people recalled, the more negative expectancies of robots becoming a threat to humans did they report. People's technical affinity and locus of control, however, appeared to diminish those fears.

3.2. *Research Paper 2:* Expectations vs. Actual Behavior of a Social Robot: An Experimental Investigation of the Effects of a cial t's nteracti n ill Level and Its Expected Future Role on e le 's al ati ns (Horstmann & Krämer, 2020)

To further examine what affects people's perceptions and evaluations of social robots, this study focuses on three potential influences: 1) the robot's behavior in form of a high or low level of interaction skills, 2) the robot's expected future role as helpful assistant or threatening competitor, and 3) people's individual backgrounds regarding robots and technology in general. When interacting with artificial entities, the behavior appears to be of substantial importance and to exert great influence (Rickenberg & Reeves, 2000). Consequently, a social robot's level of interaction skills should affect people's evaluations tremendously, in particular since interaction skills are essential for a social robot to serve its purpose (Dautenhahn, 2007; Davis, 1993). With regard to expectations, the preceding study (study II, research paper I) confirmed that most people have not personally interacted with a social robot yet and thus have to acquire information from other sources, which are often mass media. However, fictional robots are portrayed as both, "good" as well as "bad" characters, which leads recipients to experience double-minded feelings (Bruckenberger et al., 2013). Since expectations work as perceptual filters of the reality (Burgoon, 1993; Burgoon & Le Poire, 1993), expecting a social robot to either become a helpful assistant or a threatening competitor should affect people's perception of this robot, particularly in combination with low or high interaction skills. As the previous studies (study I and II, research paper I) indicated, the individual background of the person interacting with an artificial entity needs be considered as well when examining human-robot interaction processes. Therefore, this study further includes people's technological expertise, their previous experiences with real or fictional robots, and their negative expectancies regarding robots in the analyses.

Against this background, an experimental lab study with a 2x2 between-subjects design and 162 participants was conducted. Before people interacted with the social robot Nao, they were told about the motivation for having the robot interact with different humans. Respectively one half of the participants were told that the robot interacts with people to become a better, more helpful assistant, which will be able to relieve humans from onerous and exhausting tasks one day. The other half was told that the robot's goal is to learn from humans in order to become better, e.g., more efficient than them and to take over tasks one day which are currently executed by humans. During the subsequent interaction with Nao (see Figure 3), the robot presented either high or low interaction skills. High interaction skills comprised a natural dialogue (via wizard of oz design; see Dahlbäck et al., 1993), while low interactions skills included misunderstandings, mispronunciations, and restricted answering options. Participants' individual backgrounds as well as their evaluations of the robot and the interaction with it were assessed via questionnaires.

Figure 3. Experimental setting of study III: interaction with the social robot Nao



A structural equation model was built to get a comprehensive picture of the different influences of and relationships between the robot's behavior, the user's expectation, and the user's individual background. Based on the results, the robot's interaction skill level, i.e., the robot's behavior, had the strongest effect on people's evaluations. More specifically, poor interaction skills elicited a more negative evaluation than high interaction skills. Regarding the robot's expected future role, the robot was evaluated as less sociable when it was described to become a competitor compared to an assistant. People's individual background appears to have no substantial influence on how they perceived and evaluated the robot. In conclusion, people's perception of an artificial interaction partner is foremost affected by its behavior and to a lesser extent by people's expectations, while individual backgrounds appear to play a negligible role.

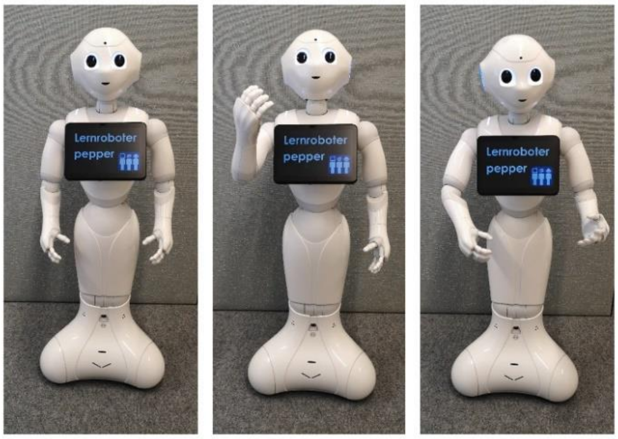
3.3. *Research Paper 3:* The Fundamental Attribution Error in Human-Robot Interaction: An Experimental Investigation on Attributing Responsibility to a Social Robot for Its Pre-Programmed Behavior

To consider attribution processes as further influencing factor for the perception of artificial entities, this study's aim is to explore how autonomous in contrast to pre-programmed behavior of an artificial entity is perceived which furthermore has an either positive or negative valence. Since human-robot interactions are becoming increasingly socially situated and multi-faceted (Young et al., 2011), the question arises whether and to what extent people perceive a social robot to be responsible for its actions and subsequent consequences. The well-known attribution theory describes how people constantly try to comprehend causes and implications of others' behaviors (Kelley, 1973; Ross, 1977). In this process, however, people appear to overestimate the influence of dispositional factors while underestimating situational ones (Ross, 1977). Transferring this to human-robot interaction, the question arises how people evaluate internal and external causes when interacting with a social robot, particularly when it unexpectedly gives unpleasant, negative feedback. This study examined whether presenting external justifications affect the perception of an artificial entity, more specifically whether a robot is evaluated more negatively when it is believed to generate impolite feedback autonomously by itself compared to when the feedback is believed to be fix-programmed in advance.

Furthermore, expectations are considered as well since study III (research paper II) indicated that the evaluation of a social robot is partly affected by previously describing the robot either as threatening competitor or as helpful assistant. These future roles are common portrayals in mass media (Bruckenberger et al., 2013; Scopelliti et al., 2005) and further reflect wide-spread expectations (Ray et al., 2008; Weiss et al., 2011). Summing up, this study's aim was to examine whether and how attribution processes occur during interactions with social robots. For this purpose, it was analyzed how a social robot's feedback is perceived when it is either described to be generated autonomously or pre-programmed while considering the valence of the feedback and the robot's expected future role.

An experimental online study with a 2x2x2 between-subjects design was conducted with a final sample of 394 participants. First, a vignette was presented in which the social robot Pepper was either described as assistant or competitor and its feedback to be generated autonomously by the robot or to be pre-programmed by programmers. After watching a tutorial video, participants took a quiz executed by Pepper during which the robot provided either constantly positive or negative feedback regarding the participants' answers (see Figure 4).

Figure 4. Experimental setting of study IV: interaction with the social robot Pepper



Based on the results, a social robot is perceived to possess more agency, to be more responsible for the feedback content, and to be more competent when it is believed to give feedback autonomously compared to when the feedback is believed to be pre-programmed. Furthermore, the more agency a robot is perceived to have, the better it is evaluated, particularly its sociability and the interaction with it. However, the robot's autonomy merely had an indirect effect via perceived agency on the evaluation of the robot. The valence of the feedback alone affected the evaluation directly and without any influence by how the feedback was generated. These results indicate that people neglected external causes for the robot's behavior, i.e., its programming, which points to the occurrence of the fundamental attribution error (Ross, 1977).

3.4. *Research Paper 4:* I Just Wanna Blame Somebody, Not Something! Reactions to a Computer Agent Giving Negative Feedback Based on the Instructions of a Person

To extend the insights regarding attribution processes in interactions with artificial entities, this study further examines how different types of agency (human-controlled vs. computer-controlled vs. mix of human- and computer-controlled) and different levels of expertise (expert tutor vs. peer study partner) influence the perception of an interaction partner who provides exclusively negative feedback. Previous research primarily focused on interaction partners which are clearly controlled either by a person or a computer program (avatar vs. virtual agent; e.g., Blascovich et al., 2002; Krämer et al., 2017; von der Pütten et al., 2010). Since interaction technologies are becoming increasingly reliable, more tasks are delegated to autonomous machines (Gogoll & Uhl, 2018; Mosier et al., 1997). Against this background, a mixture of human- and computer-controlled agent is emerging: a computer agent that represents persons in certain situations by acting on their instructions called agent representative (de Melo et al., 2018). Accordingly, since the borders between the categories person and computer increasingly blur, this study examines to whom or what responsibility for an artificial entity's behavior is attributed to. When interacting with an agent representative, the agent's behavior may be attributed "externally" to the person who instructed it or "internally" to the computer program itself (cf. attribution theory; Kelley, 1973), or in parts to both. To investigate the attribution processes in interactions with artificial entities, this study examines whether there are differences in people's reactions to negative feedback when they believe to be either interacting with a person's avatar, a computer-controlled agent, or an agent representative constituting a hybrid form of the former two. Additionally, the expertise of the interaction partner is considered since it should also affect how people react to feedback (Bannister, 1986). A peer study partner with similar expertise should be perceived as ally (Nass et al., 1996; Wilson et al., 1965), while expert tutors are generally expected to be authoritarian as well as credible (Berlo et al., 1969; Hovland et al., 1953).

In an experimental lab study with a 2x3 between-subjects design, 195 people were told to either study together with a person-controlled avatar, a computer-controlled virtual agent, or a computer-controlled and person-instructed agent representative (see Figure 5). To manipulate the expertise of the interaction partner, participants were further told to either interact with an expert tutor or a non-expert peer study partner. During the joint study time, the interaction partner would always give negative feedback regarding the participant's performance. Afterwards, people reported on their current affect and evaluated their interaction partner as well as the interaction with it.

Figure 5. Experimental setting of study V: interaction with the virtual agent Brad (left: study partner with low level of expertise, right: tutor with high level of expertise)

Although the negative feedback apparently led people to evaluate their interaction partner poorly in all conditions, there were still some small, interesting differences regarding the type of agency. Perceived agency was highest for the person-controlled avatar and similarly low for the agent representative and the virtual agent. Attribution of blame to the person(s) behind the software and negative affect was highest for the avatar and lowest for the virtual agent. No significant difference regarding negative affect and attribution of blame was found between agent representative and avatar as well as between agent representative and virtual agent. Descriptive values further indicate that the agent representative's evaluations lay between the evaluations of the avatar and the virtual agent with no significant difference to either one of them. Level of expertise had no significant effect and other evaluation measures were not affected.

4. DISCUSSION

The focus of the five studies presented in this cumulus was to examine which factors have which influence on the perception of artificial entities, individually and in interaction with each other. How people perceive artificial entities is highly relevant since it may affect how they are evaluated and treated, how people react to them, and whether they are accepted and used in the long run (Dautenhahn, 2007; Davis, 1989, 1993). Furthermore, learning about how artificial entities are perceived in comparison to how other persons are perceived will provide us with a more in-depth understanding of human perception in general. Perception particularly plays a role for artificial entities in comparison to other technologies, since they are non-living, electronic devices or computer programs but appear to be somehow alive due to their social and life-like behavior (Bartneck, Kanda, et al., 2007; Dautenhahn, 2007; Reeves & Nass, 1996; von der Pütten et al., 2010). Thus, by examining how perceptions are formed in interactions with artificial entities we are able to make better predictions for how people will feel and behave in future scenarios where these technologies are used more extensively. In the following, the results of the five studies of the cumulus are discussed with regard to the overarching question which factors influence the perception of artificial entities to which extent and how do these factors function in relation to each other. This is followed by a derivation of theoretical as well as practical implications for the field of human-computer interaction and an overall conclusion at the end.

4.1. Influencing Factors for the Perception of Artificial Entities

To examine how individuals perceive artificial entities, various factors need to be considered individually as well as in interaction with each other. First, a closer look was taken at people's expectations regarding artificial entities, whereby mass media particularly science fiction representations appeared to play a major role (study I and II, research paper I). Specifically, two contradicting but similarly prominent views on the future role of artificial entities were considered throughout the studies: artificial entities either becoming competitors or assistants to humans (study III, research paper II; study IV, research paper III). Another aspect which was included in the studies as well is people's technological expertise, more specifically technical affinity and locus of control, which also can have an effect on the formation of expectations (study II; study III). Moreover, attribution processes were taken into account in order to determine whether an artificial entity's actions are perceived to be caused internally or externally and whether this affects how the entity is perceived (study IV; study V, research paper IV). The last influencing factor included was the artificial entity's behavior, which was considered in different forms (level of interaction skills in study III; valence of feedback in study IV). The results regarding these influencing factors are outlined and discussed below.

4.1.1. Expectations

There were two predominant reasons for exploring the influencing power of people's expectations on their perception of artificial entities. First, expectations often work as perceptual filters (Burgoon, 1993; Burgoon & Le Poire, 1993) and thus may alter how we perceive an artificial entity during an actual encounter. Second, due to a lack of personal experiences with artificial entities, expectations concerning them appear to be heavily influenced by mass media portrayals, which suggest that artificial entities are far more advanced than it is actually the case (Banks, 2020; Bruckenberger et al., 2013; Sandoval et al., 2014). These inflated expectations may lead to expectancy violations when confronted with reality (Komatsu & Yamada, 2011; Kwon et al., 2016), which should have a negative effect on people's perception of artificial entities (Burgoon & Hale, 1988). Moreover, two contradicting views on social entities are prevalently promoted in the media – one predicting artificial entities to be desirable, helpful assistants and the other drawing on people's fear to compete with and possibly be replaced by machines (Bruckenberger et al., 2013; Ray et al., 2008; Scopelliti et al., 2005; Weiss et al., 2011). These contradicting views could respectively have an impact on people's mindset when encountering artificial entities and may persistently influence people's perception of them. In the following, the influence of science fiction portrayals and the two most prominent but contradicting views on artificial entities are discussed in detail.

4.1.1.1. Science Fiction as Source of Expectations

In general, it has been observed that media have an effect on people's perception of the reality as described extensively by the cultivation theory (Gerbner & Gross, 1976). Particularly, media are able to shape perceptions of technologies (Nisbet et al., 2002), which is why it is pivotal to consider prominent media representations of artificial entities when examining what influences how they are perceived. There is a large number of science fiction characters which represent artificial entities of various forms. Most of them are robots such as the Terminator, WALL-E, or R2-D2. However, there are also other forms such as the Machines from the Matrix movies, the board computer HAL 9000, or the omnipresent artificial intelligence J.A.R.V.I.S. known from the Marvel Universe. As the first two studies (study I and II, research paper I) confirm, those science fiction characters are wide-spread and well-known. As it was the case in studies conducted by Ray et al. (2008) and Scopelliti et al. (2005), the interviewees of study I also frequently referred to science fiction when describing their preferences and expectations. Here, social robots are expected and preferred to have no own emotions or an own will and not to take over tasks which involve a social aspect such as child or elderly care. Instead, robots are predominantly expected and desired to function as assistants for household- or work-related tasks. Extending these insights, study II showed that people's expectations are greatly influenced by their experiences with robots, which appear to be mainly encounters via mass media in general and science fiction in particular. Taking a closer look, these experiences appear to cause people to have rather high expectations regarding the actual abilities of social robots. Indirectly, these experiences also lead people to rather expect social robots to play significant roles in their personal future lives as well as in society in general. All these heightened expectations may be the consequence resulting from accordingly biased media representations (Bruckenberger et al., 2013; Sandoval et al., 2014).

The second study (study II) further reflected upon the influence of 25 specific fictional robot characters, which mostly could be classified as either positively (e.g., WALL-E) or negatively (e.g., the Terminator) perceived characters, on people's fears and negative expectations. Results showed, the more negatively perceived fictional robot characters people recalled, the more negative expectancies regarding social robots they reported. Those negative expectancies reflected common science fiction scenarios of humans being outraced and battled by their inventions (Khan, 1998; Ray et al., 2008; Weiss et al., 2011). With these findings regarding the influence of mass media representations on people's expectations in mind, it became clear that the influence of the two views which are promoted the most in media and portray artificial entities either in a positive, pleasant or negative, repelling way (Scopelliti et al., 2005) need to be looked at more closely.

4.1.1.2. Expected Future Role: Competitor Versus Assistant

Previous research already described how there are two prevalent views on artificial entities, particularly social robots, which are on opposite ends (Bruckenberger et al., 2013; Scopelliti et al., 2005). As Scopelliti et al. (2005) put it: "Science fiction doubtless influences people's visualization of humanlike robots: it gives a twofold image of robots: (1) as helping assistants and (2) as more frightening potential competitors or, even worse, overwhelming entities." (p. 148). Against this background, it makes sense that participants of study I (research paper I) reported that they would rather have robots take over assistive tasks, mostly related to household chores, but would feel uncomfortable with them becoming too human-like and playing a part in their social lives. Moreover, one of the main realizations of the first research paper was that mass media representations, particularly science fiction as well as fictional characters based on artificial entities, have a considerable impact on people's expectations, which should also ultimately affect their general perception of artificial entities. As a consequence, the influence of these two contradicting views were considered in the subsequent two studies.

In study III (research paper II) as well as in study IV (research paper III), people interacted with a social robot (Nao or Pepper) and were either told beforehand that the robot strives to become a helpful assistant or to take away jobs from and compete with humans. The first outlook was supposed to stress the positive assistant-expectation while the other was supposed to emphasize the negative competitor-expectation. In study III, the robot expected to become a helpful assistant was evaluated more positively with regard to its sociability compared to the robot expected to become a threatening competitor. This indicates that expectations are able to influence the perception of an artificial entity. However, no effect of this expectation manipulation was found in study IV.

One reason may be that people's expectations are formed by several aspects such as different sources of information, personal experiences, and people's technological expertise. The one-time manipulations via a short vignette may not have been sufficient to elicit extensive influences on people's perception. However, expectations should still be considered as relevant influencing factor as previous research (Komatsu & Yamada, 2011; Kriz et al., 2010; Kwon et al., 2016; Sandoval et al., 2014; Scopelliti et al., 2005) as well as the results of the first three studies (study I; study II; study III) indicate. In general, expectations should be viewed as one gear wheel in the machine that is responsible for creating people's perception of artificial entities. Another one of those gear wheels appears to be the technological expertise of the user as outlined in the next paragraph.

4.1.2. Technological Expertise of the User

How people perceive and behave around technological devices and applications depends on the design of the technology, but also on characteristics on side of the user (Lane & Lyle, 2011; MacDorman & Entezari, 2015; Rosenthal-von der Pütten & Weiss, 2015). Here, Lane and Lyle (2011) emphasize the importance of recognizing the range of technological expertise that exists among users and how this can pose actual or perceived barriers and consequently impact the adoption of new technologies. People's technological expertise may cover several facets, of which technical affinity and locus of control where considered for the studies presented in this cumulus. Technical affinity rather describes how much enthusiasm and joy a person experiences when encountering technology, particularly new, complex and exciting technology, and how this results in pronounced openness to try these technologies (Franke et al., 2019; Karrer et al., 2009; Morahan-Martin & Schumacher, 2007). Locus of control on the other hand rather reflects a person's actual competence handling electronic devices which appears to result in a more realistic view on technology (Beier, 1999; Gaul et al., 2010).

In the second study (study II, research paper I), both forms were included as potential factors for influencing people's expectations regarding artificial entities. And indeed, people with a high technical affinity had greater expectations regarding the skills of robots as well as their future role in society and their personal lives. People's enthusiasm about technological novelties appeared to have positively biased their perception of the technology. This bias apparently causes them to expect sophisticated abilities and functions, likely because this would make the technology more interesting and exciting. Moreover, people with a pronounced technical affinity also reported less fears related to robots in comparison to people with low technical affinity. Accordingly, technical affinity appears to impact people's perception of artificial entities by means of higher expectations as well as less anxiety.

Looking at another facet of technological expertise, people with a high locus of control using technology appeared to have a more realistic view on artificial entities regarding their skills as well as their future role. This can be derived from the insights of study II which show that a high locus of control was related to less pronounced expectations regarding the skills as well as future role of social robots. Moreover, locus of control also reduced the extent to which people reported fears regarding robots (on a marginally significant level), which may also be a product of the more realistic view on artificial entities. Overall, high locus of control when handling technology appears to be accompanied by a general understanding of technological devices (Beier, 1999; Gaul et al., 2010). This may enable people to make a more realistic assessment of artificial entities' capabilities, which also appears to reduce their general fears regarding non-human interaction partners.

The mitigating effects of people's technological expertise, which comprises technical affinity as well as locus of control, on people's negative expectancies regarding artificial entities were also apparent in study III (research paper II). Here, technological expertise also enhanced people's tendency to generally evaluate the interaction with an artificial entity better (on a marginally significant level), which may be a side product of generally reduced fears. In the subsequent studies (study IV, research paper III; study V, research paper IV), however, technological expertise appeared to play no particularly relevant role. Coming to a conclusion here, people's technological expertise can affect how people perceive artificial entities by influencing the formation of expectations and by reducing people's fears. However, those effects are not too pronounced and not persistent in all contexts. Thus, technological expertise should be considered as a potential influencing factor, but not as the most central one. As further factor influencing people's perceptions of artificial entities, attribution processes need to be considered, i.e., whether agency and responsibility for a behavior are ascribed to the artificial entity or whether external factors are considered to be the decisive cause.

4.1.3. Attribution Processes

As outlined in the theory section, there is a fundamental human need to reduce uncertainty (Berger & Calabrese, 1975), which is why people always make an effort to figure out who or what caused a person's behavior. This results in attributing an opponent's behavior either internally to the opponent's disposition or externally to situational, relational, or other contextual factors (Kelley, 1973). In case of behavior that is believed to be intentional, perceived motives are also considered to interpret and evaluate the behavior. In order to attribute the cause of behavior internally to an entity, it needs to be perceived to possess agency first. Agency can be defined in several ways, but basically describes the capacity to act independently and free of choice (Banks, 2019; Barker, 2005).

For the perception of agency, the focus in this work was laid on perceived autonomy and intentionality as prerequisites, which together with responsibility based on situational and role-related aspects is believed to determine the internal attribution of behaviors and subsequent consequences to an artificial entity (Kelley, 1973; Shaver, 1985; Sullins, 2006). Autonomy in particular was the focus of study IV (research paper III), where the effect of autonomous compared to pre-programmed feedback on people's perception of a social robot's actions were examined in an experimental setting. Study V (research paper IV) addressed the question whether responsibility for the behavior of an artificial entity is attributed differently when it is controlled by a computer, a person, or a mix of the former two. The perception of agency was the focus of both studies by varying the extent to which the artificial entity could be assumed to act with autonomy and intentionality. Ultimately, it was examined in both studies whether varied circumstances may affect attribution processes, whereby common machine-related circumstances where emphasized (fix programming in study IV and control exerted by a program, person, or both in study V). The findings regarding the influence of agency, particularly autonomy and intentionality, as well as responsibility on attribution processes in human-machine interaction are discussed in the following alongside with resulting effects for the perception of artificial entities.

4.1.3.1. Autonomous Versus Pre-Programmed Behavior

Autonomy was described by Deci and Ryan (1987) as “an inner endorsement of one's actions, the sense that they emanate from oneself and are one's own” (p. 1025). In other words, autonomous behavior is self-governed, voluntary, and not executed on behalf of someone else (Banks, 2019). However, it might be influenced by external factors such as social, technological, and systemic forces (Rose & Truex, 2000; Sullins, 2006). The extent to which an artificial entity’s behavior can be classified as autonomous is debatable, however, the impression of acting autonomously might be decisive for how it is perceived. Against this background, study IV (research paper III) showed how autonomous behavior by an artificial entity leads people to perceive the entity to possess more agency compared to when the artificial entity is believed to only execute previously programmed actions. Along with the increased perception of agency, people were also observed to attribute more responsibility and more competence to the artificial entity. Furthermore, an increase in perceived agency also leads to a more positive evaluation of the artificial entity’s sociability as well as the interaction with it. These results suggest that telling people whether an artificial entity’s actions are autonomously generated or determined by a strict program is enough to influence their perception of this entity with regard to agency, responsibility, and competence. Considering the insights of the attribution theory (Kelley, 1973; Kelley & Michela, 1980), an explanation for these findings may be that in case of autonomous behavior causes were evaluated to be internal, i.e., the robot itself is seen as responsible for its behavior. In case of the pre-programmed behavior, the cause then may be evaluated as external, i.e., the program or programmer is held responsible.

However, an important finding of study IV is that the manipulated autonomy neither had a direct effect, nor an interaction effect with the valence of the feedback on the evaluation of the robot’s sociability and the general interaction. Since a strict program which determines an artificial entity’s actions should pose a clearly external cause, the evaluation of the artificial entity was assumed to be affected less by the valence of the feedback it gives in case of pre-programmed feedback. When an artificial entity is beliefed to be able to generate feedback autonomously, the valence of this feedback was assumed to have a greater effect on how the entity is perceived. However, this was not the case, which can have several reasons. Looking at known processes from social psychology, as recommended by various researchers (Burgoon et al., 2016; Krämer et al., 2012; Nass & Moon, 2000; Reeves & Nass, 1996), the results potentially point to the occurrence of the fundamental attribution error (Ross, 1977). The fundamental attribution error describes how despite logical explanations pointing to external factors determining a person's behavior, the person's disposition is considered as more decisive (Jones & Harris, 1967; Ross, 1977). In other words, internal factors are overestimated, while external factors are neglected (Ross, 1977). Transferring this to the setting of this study, the artificial entity itself might have been perceived to be responsible for the feedback because the program was neglected as external cause.

Another explanation could be that artificial entities are generally perceived as not able to act completely autonomously, at least not based on own intentions. Considering the results of the first study (study I, research paper I), people do not expect artificial entities to possess own intentions, which might lead them to perceive artificial entities' general scope of actions to be limited (Bigman & Gray, 2018; Malle & Knobe, 1997). Furthermore, artificial entities might be perceived to underlie general technological restrictions, which are not present with human interaction partners (Sullins, 2006). However, if the robot would not have been hold accountable for its actions in general, the valence of the feedback should not have had any effect on the evaluation of the robot, which was not the case. Also, the results regarding perceived agency, competence, and responsibility for the feedback further contradict this explanation.

As the results of the media equation studies point out, people react automatically and mindlessly to artificial entities (Nass & Moon, 2000). Logic explanations for an artificial entity's behavior appear to be neglected here as well, otherwise people would not respond positively to flattery by a computer (Fogg & Nass, 1997b) or reciprocally help a computer that helped them before (Fogg & Nass, 1997a). Research was conducted specifically to examine whether people think of the programmer when interacting with a computer, which appeared not to be the case (Nass & Moon, 2000; Sundar & Nass, 2000). Coming to a conclusion, the absence of an interaction effect between the robot's autonomy and the valence of its behavior indicate that external factors controlling an artificial entity's behavior are neglected. The meaning of the behavior alone appears to be decisive for subsequent overall evaluations of an artificial entity. This finding can be explained as some form of attribution error (Ross, 1977) and as disregard of the program's and/or programmer's influence (Nass & Moon, 2000). Both explanations basically amount to the same thing: an overestimation of factors related to the artificial entity and an underestimation of external factors controlling, determining, or limiting the artificial entity's actions.

Summing up, an artificial entity's autonomy has an impact on the perception of some of its features such as agency, responsibility, and competence, its general evaluation however appears to be determined predominantly by its behavior. But do attribution processes differ if the behavior by an artificial entity is described to be controlled by a computer program, a person, or a mix of the two? To address this question in depth, study V (research paper IV) compares three different forms of agency, which may exert control over an artificial entity's behavior: a person-controlled avatar, a computer-controlled virtual agent, and a hybrid of the former two.

4.1.3.2. Person Versus Computer Versus Hybrid Agency

Since the recent evolvement of interactive technologies makes it progressively feasible to use artificial entities as convincing and reliable representatives (e.g., Google duplex; Leviathan & Matias, 2018), the question emerges how the actions of these technological representatives are perceived and to whom responsibility for the actions is attributed to. Three different forms of agency were compared, an avatar controlled by a person in real-time, a virtual agent controlled by an autonomous computer program, and a hybrid agent representative, which is an autonomous computer program that acts upon specific instructions given by a person to represent this person in a certain situation.

Confirming previous assumptions (Blascovich, 2002), people reported to perceive more agency when interacting with a person's avatar compared to the virtual agent and the agent representative. Both agent forms were evaluated to possess an equally low amount of agency, which can be explained by the fact that both were restricted in their freedom to act, by a program and/or a person's instructions. The avatar on the other side was described to the participants as only being the graphic representation of a person and that they would interact live with this person only via the avatar. Thus, the avatar was probably equated with the person behind, which would explain the high amount of perceived agency in those conditions.

Since all forms of interaction partner displayed rude, mean behavior directed towards the participants, attribution of blame and the influence on people's mood were of particular interest. Results showed that person(s) behind the software were blamed the most in case of an avatar and the least in case of the virtual agent (marginally significant difference). This further confirms that in case of an avatar, people were aware of interacting with a person, which was not the case with a virtual agent. Interestingly, with the agent representative, which represents a mixture form, the amount of blame attributed to the person(s) behind the software was lower than with an avatar, but higher than with a virtual agent (with no significant difference to either one of the other forms). In a similar vein, the most negative affect was reported after people interacted with a rude avatar and the least after interacting with a virtual agent. This may be explained by people's need to belong which is accompanied by a strong aversion to rejection by other people (Baumeister & Leary, 1995). Interacting with an avatar, the rejection must have been perceived to come more directly form a person resulting in more negative affect compared to a virtual agent or an agent representative. Looking at the agent representative specifically, values for negative affect were located again between the two other forms (with a marginally significant difference to the avatar and no significant difference to the virtual agent).

Altogether, the results indicate that an agent representative, in a subtle way, is perceived to be a mix of a computer-controlled and person-instructed entity, which elicits reactions that are somehow similar to both types as if it forms a mean of the two. Furthermore, with an agent representative, attribution of responsibility appears to be challenged and blame for rude behavior to be deflected (Bivins, 2006; Royzman & Baron, 2002). Of course, these results have to be interpreted with caution given the small, partly not significant differences. However, the observed tendencies suggest slight but conclusive differences in the perception of artificial entities regarding their type of agency, i.e., whether they are controlled by a person, a computer program, or a mix of the former two.

Summing up, attribution processes, which are influenced by the perception of agency, particularly autonomy and intentionality, as well as responsibility, are an important mechanism to be considered as influencing factor for the perception of artificial entities. The extent to which an artificial entity's behavior is perceived to be programmed or generated autonomously and the extent to which an artificial entity is perceived to be controlled by a person, a computer, or a mixture of the two may particularly affect the perception of agency, although just in a subtle way. Pervading all studies, the behavior appeared to be attributed internally to the artificial entity which may be explained by the fundamental attribution error (Ross, 1977) and underlines the relevance of the actual behavior during an interaction (Rickenberg & Reeves, 2000). In general, expectations and attributions appear to have effects which are limited to certain aspects or characteristics of an artificial entity. Behavior, particularly with a clearly positive or negative valence, though appears to have an overarching influence on people's perception with strong and pervasive effects, which is elaborated in the following.

4.1.4. Behavior

According to Rickenberg and Reeves (2000), an artificial entity's behavior has a major impact on how it is evaluated subsequently: "the ultimate evaluation is similar to those for real people - it depends on what the character does, what it says, and how it presents itself" (p. 55). As outlined at length before, behavior can have many forms and covers many aspects of interactions with artificial entities. Along with the variety of possibilities to alter behavioral facets of artificial entities, there is an extensive body of research studying how these alterations affect people's perception, evaluation, and acceptance of as well as their reaction to non-human interaction partners (cf. Bergmann et al., 2012; Eyssel et al., 2012; Häring et al., 2011; Horstmann et al., 2018; Mumm & Mutlu, 2011). In the studies presented in this cumulus, the focus was laid on actual interactions with artificial entities, whereby the role of behavior clearly came to the fore. More specifically, how low versus high interaction skills as well as negative versus positive feedback influences perceptions and evaluations were explored during those interactions. The respective results are discussed in the following.

4.1.4.1. Level of Skills

Particularly interaction skills are essential for artificial entities since their fields of application usually involve social encounters with humans, for instance, as household and health care assistant, entertainment, and information guide (Dautenhahn, 2007). According to Dautenhahn (2007), the ability to communicate with high-level dialogue or at least to give the impression to do so, resembles the core definition of a social robot. Against this background, in study III (research paper II) people interacted with a robot displaying either sophisticated communication skills allowing a natural interaction or restricted, poor communication skills including limited answer options, misunderstandings, and false grammar. Confirming previous assumptions, a robot in possession of highly developed interaction skills was evaluated better with regard to its sociability as well as competence than a poorly skilled robot. Likewise, the interaction in general was also evaluated better in case of a highly skilled robot, which is probably because the interaction was more natural, took less effort, and was thus more pleasant. Furthermore, artificial entities might be expected to possess high interaction skills in order to fulfill their tasks which particularly involve interacting socially with people (Dautenhahn, 2007). With poor interaction skills, the usefulness of these technologies was possibly perceived to be very low and low perceived usefulness of technologies has generally been observed to elicit negative attitudes as well as low usage intentions (Davis, 1989, 1993). Moreover, a social robot displaying poor interaction skills may also have caused a negative expectancy violation.

Since expectancy violations were found to be related to detrimental communication outcomes (Burgoon & Hale, 1988), this might also be an explanation for the negative evaluation of the robot displaying low interactions skills. How the robot's behavior was perceived was neither affected by people's technological expertise, nor by their experiences with robots, nor by previously describing the robot's future application to be an assistant or a competitor. Thus, the behavioral influence appears to be so strong that it affects all evaluation measures significantly and is not affected by other influencing factors.

4.1.4.2. Valence of Behavior

Most artificial entities fill out the role of a service provider, for instance, when interacting with customers, clients, or patients. Because of that role and their general human-likeness, artificial entities are expected to behave in accordance with social norms, particular with regard to politeness (Sayin & Krishna, 2019). Thus, one of the research objects of study IV (research paper III) was to examine how negatively or positively valenced feedback behavior affects people's perception of a social robot. As it was presumed, negative, impolite feedback led to a substantially more negative evaluation of the robot's sociability as well as the interaction in general compared to positive, benign feedback. Besides these rather predictable results, it became also apparent that the effect of the behavior is not affected by a manipulation of its cause. In other words, whether people were told that the feedback was generated autonomously by the robot itself or fix-programmed in advance did not alter how the either benign or offensive behavior influenced the evaluation of the robot and the interaction. Likewise, in the last study (study V, research paper IV), the virtual agent was also evaluated very poorly after presenting negative feedback which was neither affected by information regarding the type of agency controlling the virtual agent (computer program, person, or hybrid of the former two), nor by its level of expertise. Thus, again the effect of the behavior appears to be so strong that other influences seem to subordinate, which is illuminated in more detail in the following.

4.2. Relations Between Influencing Factors

After reflecting upon the main effects of the different influencing factors on the perception of artificial entities, we are going to take a closer look at the relationships between those factors. Here, it is evaluated how the identified factors influence and function in interaction with each other. Particularly, the following section reflects upon a) expectations and technological expertise functioning as filters for the perception of artificial entities, b) the programming of artificial entities as external cause as well as determining force for their behavior, and

c) artificial entities' behavior as strong, pervasive, and overarching influence on people's perception of artificial entities.

4.2.1. Expectations and Technical Affinity as Perceptual Filters

According to Goffman (1974), people's experience and understanding of the real world is influenced by their primary framework, which gives meaning to otherwise meaningless aspects of a situation. More specifically, framing influences how people interpret and process information (Goffman, 1974) and expectations may work as framing devices for these processes (Burgoon, 1993), as well as people's previous experiences and general knowledge. As Burgoon (1993) explains, expectancies may "serve as perceptual filters, significantly influencing how social information is processed" (p. 32). This was shown in a series of experiments where positive or negative expectations regarding an interaction partner were induced before an encounter. In contrast to negative induced expectations, positive induced expectations led to a more positive evaluation of the interactions partner and the partner's communication behavior (Burgoon & Le Poire, 1993; Honeycutt, 1991; Ickes et al., 1982). The authors explain that expectations "persevere and influence subsequent information processing, behavior, and perceptions" (Burgoon & Le Poire, 1993; p. 68).

Based on the media equation theory (Reeves & Nass, 1996) as well as the overview of differences and similarities between human-human and human-machine interaction presented by Krämer et al. (2012), these processes should also take place in human-machine interactions. Thus, expectations regarding artificial entities should generally also shape how they are perceived subsequently. This was found to be the case for some characteristics of the artificial entity in the presented studies, particularly with regard to the perception of sociability, competence, responsibility, and agency. Expecting an artificial entity to pursuit the role of a helpful assistant enhanced its perceived sociability (study III, research paper II), expecting it to act autonomously intensified how competent and responsible it was perceived (study IV, research paper III), and expecting the behavior to be restricted by external forces reduced the amount of perceived agency (study IV, research paper III; study V, research paper IV).

The sources and mechanisms for developing and adapting expectations for artificial entities are manifold. Besides experiences, personal or through mass media, people's expertise regarding technology in general has a considerable impact (study II, research paper I). An increase in technological expertise changes people's attitudes towards a technology, for instance, regarding possible barriers and concerns (Berge et al., 2002). People who feel intimidated by technology may experience a loss of control and feel uncomfortable when using technology (Parisot, 1997), which should ultimately lead to a more negative perceptions of technologies, particularly new, unfamiliar ones. Higher skills and knowledge regarding technologies on the other hand enhances people's ability to cope with new technologies, i.e., whether they are actively approached or rather avoided (Franke et al., 2019).

Technological expertise has many facets of which two were looked at more closely: people's enthusiasm described as technical affinity (Karrer et al., 2009) and their locus of control based on actual competence, knowledge, and skills (Beier, 1999). Both aspects are highly connected, however, there are some small changes regarding their exact effects. Results of study II (research paper I) show that technical affinity leads to inflated expectations, while locus of control predicts lower, more realistic expectations. Both facets of technological expertise appear to result in less fears and negative expectations regarding artificial entities (study II) and a more positive evaluation of the interaction with it (marginally significant effect; study III, research paper II). Thus, technological expertise appears to shape people's expectations regarding artificial entities in an either realistic or enhanced way, which than may has an effect on people's general perception of them. Furthermore, people's technological expertise reduces their fears regarding artificial entities, which should ultimately also affect people's perception to a great extent (Nomura et al., 2008; Nomura et al., 2012).

Summing up, technological expertise appears to have an impact on the formation of expectations, particularly negative ones, which then may function as perceptual filters when interacting with artificial entities. Technological expertise as well es expectations were found to affect the perception of artificial entities with regard to certain characteristics. However, these effects were rather minor and non-pervasive, particularly when looked at in combination with other more prevalent factors such as behavior. Whether an artificial entity is perceived to be responsible for its behavior or whether the behavior is perceived to be determined by external factors such as programming or a person's instructions should affect how it is perceived after an interaction, which is illuminated in the following.

4.2.2. Programming as External Cause of Actions

As we learned from the attribution theory, there are internal or external causes for an agent's behavior (Kelley, 1973). Identifying the decisive factor(s) for causing a certain behavior already poses a challenge in human-human interactions, which often results in attribution errors (Nisbett & Ross, 1983; Ross, 1977). With artificial entities, however, this poses an even greater challenge since it is hard to determine how their behavior is generated, who or what is controlling it with what intentions, and who or what is ultimately responsible (Sullins, 2006). Against this background an extensive body of research deals with the topic of moral agency of artificial entities (e.g., Banks, 2019; Bigman & Gray, 2018; Gogoll & Uhl, 2018; Himma, 2009; Sullins, 2006). According to Sullins (2006), prerequisites for possessing moral agency are autonomy, intentionality, and responsibility. All these characteristics are by definition hard to ascribe to artificial entities, for instance, since they are usually dependent on some sort of programming and not able to develop intentions based on personal needs or motives (Sullins, 2006). However, since human traits were found to be attributed to artificial entities (e.g., Eyssel & Hegel, 2012; Horstmann et al., 2018; K. M. Lee et al., 2006), an effect of behavior that is perceived to be performed autonomously by an artificial entity must be considered.

Study IV (research paper III) showed that when emphasizing an artificial entity's dependence by describing its behavior as being rigidly pre-programmed, this affects how certain aspects of the entity are perceived. Particularly its agency, competence, and responsibility for the behavior are rated lower compared to an artificial entity described to act autonomously. Likewise, in study V (research paper IV), emphasizing that a computer program was controlling an agent's behavior led people to perceive less agency. Thus, on a conscious, cognitive level people appear to be aware that the program is an external force controlling the artificial entity's behavior.

Nonetheless, explicitly negative behavior was observed in both studies (study IV and V) to elicit intense (emotional) reactions, which reflected in strong negative evaluations of the artificial entity exhibiting this behavior. Therefore, on a non-cognitive, rather automatic, behavioral level, people seem to neglect the programming as external cause for an artificial entity's behavior and rather ascribe the behavior to the artificial entity itself. This goes in line with quite a few studies which showed that people overlook the artificial nature of machines, while perceiving them to possess autonomy as well as intentions and ascribing responsibility to them (Asher et al., 2012; Horstmann et al., 2018; T. Kim & Hinds, 2006; Nass & Moon, 2000). Here, the life-like behavior is seen as eliciting those reactions (Bartneck et al., 2009; Bartneck, Kanda, et al., 2007; Bartneck, van der Hoek, et al., 2007), which appears to override the cognitive evaluation of the artificial entity's programming as external cause for its behavior. Instead, the cause appears to be perceived to be located inside of the artificial entity, which goes in line with the fundamental attribution error (Ross, 1977) and leads to the necessity to consider the actual behavior as primary factor for influencing people's perceptions, as further outlined below.

4.2.3. Behavior as Overarching Influence

Reflecting upon all the studies summarized in this cumulus, an artificial entity's behavior proves to be the most pivotal factor influencing how it is perceived. Furthermore, it is remarkable that this influence seems to be unaffected by any circumstantial variations such as negative versus positive expectations or autonomy versus fixed programming. It seems like people's attention is shifted away from any surrounding information and shifted towards the behavior and what it means or rather how it makes people feel when interacting with an artificial entity. A similar explanation was also used for the effects described by the media equation theory: "cues trigger various scripts, labels, and expectations, which in turn focus attention on certain information while diverting attention away from other information." (Nass & Moon, 2000; p. 83). However, it needs to be noted that the behavior examined in the current studies always had a very clear meaning, which was either positive (high interaction skills and benign feedback) or negative (low interaction skills and offensive feedback). Particularly, the effect of expectations are believed to be higher in case of ambiguous behavior that has no clear meaning (Burgoon & Hale, 1988), which may also apply to attribution processes. Thus, the insights here are only based on easily interpretable behavior with a clearly negative or positive valence.

Summing up, based on the results of the studies of the cumulus and in line with previous research in this area (Dautenhahn, 2007; Fogg & Nass, 1997a; Nass & Moon, 2000; Rickenberg & Reeves, 2000), particularly unambiguous behavior of an artificial entity needs to be considered as most pivotal influence on its perception, as well as evaluation and acceptance. In the following chapter, theoretical implications will be discussed in terms of how the main findings of the five studies contained in this cumulus, particularly with regard to the overarching influence of the artificial entity's behavior, align with and further extend previous research.

4.3. Theoretical Implications

The insights gained from the five studies presented in this cumulus confirm and extend previous research regarding the perception processes of humans in general and for interactions with artificial entities in particular. The media equation theory describes how social cues, which are often behavioral cues, cause people to react to artificial entities as if they were alive while neglecting that they do not need to be treated this way due to their non-human, technological nature (Nass & Moon, 2000; Reeves & Nass, 1996). Research showed that people do not think of the programmer when reacting this way (Nass & Moon, 2000; Sundar & Nass, 2000) and do not recall treating an artificial entity consciously like a human interaction partner (Nass et al., 1999; Nass & Moon, 2000). Similar processes seem to occur with regard to influencing factors for the perception of artificial entities. On a cognitive level, people apparently evaluate certain aspects of an artificial entity by considering the information given to them beforehand, for instance, regarding the entity's future role or how its behavior is generated. However, overall evaluations, particularly with regard to how positive the interaction and how sociable the artificial entity are perceived, are primarily affected by its behavior. Negatively valenced behavior, for instance, in form of poor interaction skills or harsh criticism, appears to elicit particularly strong emotional reactions, which are rather mindless and, as described in the previous chapter, draw attention away from other influencing factors. In general, it can be concluded that well-tried social psychological theories may be used as valuable frameworks to understand the dynamics in interactions with artificial interaction partners as suggested by other researchers before (Burgoon et al., 2016; Krämer et al., 2012; Reeves & Nass, 1996). In the following, the theories that were considered for this thesis are summarized and it is explained how the results of the current studies deliver further theoretical insights against the background of these well-established theories.

4.3.1. Expectations

Let us begin with the theoretical implications regarding the influence of expectations. As described by the uncertainty reduction theory (Berger & Calabrese, 1975), there is a fundamental human need to diminish uncertainty regarding an unfamiliar interaction partner by gathering and inferring information about the other. The goal here is to be able to explain and predict the interaction partner's behavior, which explains the emergence of expectations. The researchers Darley and Fazio (1980) name observation of behavior (how the target acts), class affiliation (what class or category the target is assumed to belong to), and reputation (what others expect or feel about a target) as determinants for the development of expectations. In other words, expectations may stem from personal experiences, reports from others, and/or by assigning a certain category. In human-machine interaction, personal experiences are still lacking (Banks, 2020) and people have trouble assigning a clear category since artificial entities are technological devices that nonetheless appear to be somehow alive (Kahn et al., 2011). Consequently, reports by others should play a considerate role in the formation of expectations regarding artificial entities. Particularly, mass media, which are wide-spread and easily accessible, are able to fill that gap with information and to reduce people's uncertainty. Thus, the impact of expectations inspired by mass media portrayals on the perception of artificial entities needs to be looked at with great care (Banks, 2020; Nisbet et al., 2002).

The two most prevalent views on artificial entities promoted by mass media, particularly science fiction formats, are that of an useful assistant and that of an undesired competitor (Scopelliti et al., 2005). Previous research presumed a substantial influence of these two prominent views, but never systematically investigated their impact during an actual interaction with an artificial entity (Bruckenberger et al., 2013; Scopelliti et al., 2005). This was tackled by study III (research paper II), which showed that expecting an artificial entity to take over the role of an assistant apparently leads to perceive this entity to be more sociable than when it is expected to become a competitor. Based on the results it can be said that emphasizing either one of those two views – assistant versus competitor – in a previous description influences how sociable an artificial entity is perceived.

However, it needs to be noted that the overall evaluation of the interaction as well as other characteristics of the artificial entity were not affected by the previously presented description. This indicates that the perception of an artificial entity is multi-facetted and assembled based on various pieces of information. Personal experiences also determine the formation of expectations (Darley & Fazio, 1980) and while interacting with the robot, people were able to form or adapt what they think of artificial entities in general and of this robot in particular. Moreover, in the presented study only a simplified, binary version of expectations were used, while expectations in reality are rather found on a continuum and within a complex network of various norm-related and person-related expectations concerning different facets of another person (Burgoon & Hale, 1988). However, in general the results lead us to the conclusion that people's perceptions of artificial entities regarding certain characteristics are affected by their expectations, which can be manipulated or at least reinforced by means of a brief description.

4.3.2. Technological Expertise

Unlike in human-human interactions, technology-related characteristics on the user-side need to be taken into account as factors influencing the formation of expectations and ultimately the perception of artificial entities. According to Lane and Lyle (2011), "a person's use of and motivation to adopt technologies are to some extent idiosyncratic processes" (p. 40). This is underlined by researchers in the field of human-robot interaction, who argue that personality traits have an important impact on the perception of and interaction with robots (MacDorman & Entezari, 2015; Rosenthal-von der Pütten & Weiss, 2015). A study looking at the impact of user characteristics such as age, gender, and technological expertise on the use of technologies revealed a striking influence of technological expertise on perceived barriers (Lane & Lyle, 2011). Furthermore, technological expertise may reduce the amount of uncertainty which people may experience when encountering novel technologies such as artificial entities. Less uncertainty should then enable people to explain and predict the behavior of artificial entities better, which should result in less discomfort or anxiety regarding these technological interaction partners (Berger & Calabrese, 1975).

Technological expertise is a wide term and covers different aspects, which are related, but may slightly differ in their effects on people's perception of artificial entities. This work looked at technological affinity (Karrer et al., 2009) as well as locus of control (Beier, 1999). Both constructs correlated strongly and had similar effects, however two interesting observations could be made during the studies of this cumulus. First, technological affinity appeared to enhance expectations regarding robots' skills and their future role, while locus of control appeared to have a diminishing effect (study II, research paper I). Technological affinity represents enthusiasm about particularly novel technologies (Karrer et al., 2009), which apparently causes tendencies to overestimate abilities and potential functions of artificial entities in the future. Locus of control, in contrast, may be more based on actual experience and competence handling technologies (Beier, 1999), which appears to result in a more grounded, realistic assessment. Second, a tendency for stronger effects coming from technological affinity than locus of control could be observed. This further underlines the slight difference between those two constructs, since all artificial entities used in the studies were at that point novel technologies which are not encountered often in daily lives (Banks, 2020). Thus, it makes sense that enthusiasm about new technologies plays a bigger role than locus of control which is based on actual competence and experience with the still uncommon technology.

Both technology-related character traits, technological affinity as well as locus of control, were found to have a reducing effect on people's negative expectations and fears in study II. This confirms the assumption that technological expertise diminishes uncertainty regarding artificial entities. As we already know from uncertainty reduction theory (Berger & Calabrese, 1975), people experience a high need to reduce uncertainty. Fears might particularly stem from a high amount of uncertainty due to few or no personal encounters accompanied by threatening media portrayals and a lack of technological expertise to relativize those portrayals (Bruckenberger et al., 2013; Weiss et al., 2011). Against the results of the current studies, it can be strongly argued that technological expertise needs to be taken into account as important influence on the formation of expectations and ultimately the perception of artificial entities.

4.3.3. Attribution Processes

Third, attribution processes were considered as influences on the perception of artificial entities. The attribution theory describes how behavior is either attributed to internal or external causes, i.e., a person's disposition or situational factors (Kelley, 1973). In addition, in case of behavior that is perceived to be caused by intentions, motives for this behavior are attributed (Malle et al., 2001; Reeder et al., 2002). Both, the attribution of causality as well as intentionality ultimately determine how an act is interpreted and how the actor is perceived, particularly if the action is believed to be caused by the person and to be based on the person's personal motives. As prerequisites for these attribution processes research names agency, particularly autonomy and intentionality, as well as responsibility (Kelley, 1973; Malle et al., 2001; Sullins, 2006). Autonomy and the attribution of agency were particularly considered in the studies of this cumulus since previous research showed that the more control a person is perceived to have over a situation, the more responsibility is assigned to that person (Sosis, 1974).

The results of study IV (research paper III) indicate that portraying an artificial entity to be acting autonomously results in perceiving the artificial entity as being in control of its behavior, competent as well as responsible for its actions. Likewise, the results of study V (research paper IV) suggest that describing an artificial entity to be controlled by a person, computer program, or a mix of the two affects its perceived agency accordingly. Consequently, it can be ascertained that the evaluation of certain characteristics of an artificial entity is affected by the extent it is described to be acting autonomously. Furthermore, it was emphasized that the autonomously acting robot would decide on its own how to respond to the participants, which may have led participants to perceive the behavior to be intentionally as well. With a view to human perception in general, we learned that an interaction partner is evaluated differently, for instance, more responsible, when framed to be acting autonomously and intentionally.

However, it needs to be noted that pervading study IV as well as study V, the overall evaluations of the artificial entities' sociability as well as of the interactions were not directly affected by framing them with or without autonomy or by being controlled by a person, computer program, or a mix of the two. Only the valence of the artificial entities' behavior appeared to play a decisive role here. These results may speak for the overshadowing impact of the actual behavior as outlined below, but may also be a sign for the occurrence of the fundamental attribution error (Ross, 1977). During the interaction, the interaction partner and their behavior move into the focus of attention, while other factors delivering external reasons for the partners behavior are pushed into the background. Why else would a robot presenting impolite feedback be evaluated more negatively than one presenting benign feedback although it was explicitly outlined that the robot's behavior was planned in a fix way by a programmer and that the robot is not able to influence it? This phenomenon was extensively observed in human-human interactions (Jones & Harris, 1967; Jones & Nisbett, 1972; Nisbett & Ross, 1983; Ross, 1977) and now the results particularly of study IV and partly also of study V suggest that people also tend to overestimate the influence of an artificial entity's disposition while neglecting external factors causing, controlling, or influencing the entity's behavior.

Summing up, the cognitive evaluation of certain characteristics that were directly related to the description regarding the artificial entity's autonomy were affected by this description. Thus, the human perception of certain characteristics of an interaction partner can be influenced by a description suggesting an internal or external attribution as well as the potential presence of intentions. Overall perceptions, however, appear to be rather affected by behaviors which have an unambiguous and inherent meaning such as clearly negative feedback (Burgoon & Le Poire, 1993). Potentially, this behavior by itself suggests autonomy and intentionality and thus is then more relevant for the overall perception of an interaction partner than a previously given description, which is discussed in the following.

4.3.4. Behavior

Among the first, Rickenberg and Reeves (2000) pointed to the high relevance of the behavior of an artificial entity for its evaluation. In this vein, other research delivered evidence that an artificial entity is evaluated better after people interacted with it (Mirnig et al., 2012), which is interpreted by Rosenthal-von der Pütten and Weiss (2015) as that the artificial entity's behavior is more important than, for instance, the appearance. This further goes in line with insights of the expectancy violation theory (Burgoon & Hale, 1988) and parallels to the media equation theory can be drawn as well (Reeves & Nass, 1996) as outlined below.

Studies examining the effects of expectancy violations often employed behavioral manipulations which elicited strong reactions (Burgoon, 1993). In one study, negative or positive expectations regarding an interaction partner were induced beforehand, who then behaved either negatively or positively (thus in accordance or contrast to the previously given descriptions; Burgoon & Le Poire, 1993). Here, the actual communication behavior produced the strongest effects accounting for 47 to 60 % of the variance. As the authors sum up, "otherwise powerful expectancies still failed to negate the effects of actual interaction behavior." (Burgoon & Le Poire, 1993; p. 89). Thus, behavior generally is a powerful factor for influencing a person's emotions, thoughts, and behavioral reactions, particularly in case of behavior with a clearly negative valence which violates societal interaction norms (Burgoon & Le Poire, 1993).

The media equation theory describes how social cues cause people to neglect the fact that they are interacting with a technological device or application and shift their attention to norms regarding how to treat social interaction partners (Nass & Moon, 2000; Reeves & Nass, 1996). These social cues are predominantly behavioral cues such as interactivity and natural use of language (Reeves & Nass, 1996). Thus, what is observed here is basically that the behavior of the artificial entity has such a strong effect on people that it overshadows cognitions for instance regarding the artificial entity's technological nature.

This overshadowing effect of the behavior and its meaning in social interactions was also observed in the studies of this cumulus whenever an actual interaction took place (study III, research paper II; study IV, research paper III; study V, research paper IV). Altogether, expectations, technological expertise, as well as attribution of agency and responsibility appear to have limited effects on the perception of specific areas, while the overall perception of artificial entities appears to be mostly unaffected by these previously given information and character traits. The only factor which appeared to have an overall effect, with great effect sizes and not interacting with any of the other considered factors, was the artificial entity's behavior. In other words, regardless of whether a) the artificial entity's is expected to take over a pleasant or threatening role, b) the behavior is pre-programmed or autonomous, or c) a person, computer program, or a mix of the two is controlling the behavior, the valence or meaning of the behavior alone has the greatest impact on various perception measures such as sociability, competence, and interaction evaluation. Considering these findings, the question arises how they can be used in a practical manner to design interactions with artificial entities more successful and pleasant, which is discussed in the following chapter.

4.4. Practical Implications

Artificial entities are used in contexts in which interactions and/or collaborations with humans or other agents are often necessary (Dautenhahn, 2007) – for instance, for elderly care (Broekens et al., 2009; Tanaka et al., 2017), as support for autistic children (E. S. Kim et al., 2013), for education (Atkinson, 2002; Baylor et al., 2003; Paiva & Machado, 1998), as well as to provide services and information (Burgard et al., 1999; Gockley et al., 2005). As a consequence, social and interactive skills are increasingly required for artificial entities to fulfill their tasks in a satisfactory way (Dautenhahn, 2007). For this purpose, Dautenhahn (2007) describes how an artificial entity that takes over the role of an assistant needs to "identify and respond to the human's needs, … to ensure the human is satisfied and happy (with its behaviour), which implies showing behaviours towards the human that are comfortable and socially acceptable" (p. 700). Thus, designing an artificial entity and its behavior, it is crucial to consider its future human interaction partners, particularly how they will perceive and then respond to the artificial entity.

Based on the results of the studies presented in this cumulus, there are several factors to which attention should be paid to. First of all, expectations can influence the perception of an artificial entity in a certain direction. As study I and II (research paper I) showed, people already hold various expectations of artificial entities, mostly based on personal experiences, reception of mass media portrayals, and their technological expertise. In general, artificial entities are usually expected to either become a helpful assistant or threatening competitor (Scopelliti et al., 2005). Emphasizing the positive image of an assistive artificial entity prompts people to evaluate it as more socially attractive (study III, research paper II). Furthermore, a generally high technical affinity as well as locus of control reduce fears regarding artificial entities (study II; study III), resulting in a more positive evaluation of an interaction with an artificial entity (study III). Thus, creating or enhancing positive expectations regarding artificial entities as well as fostering technological expertise helps to evoke a positive perception, which ultimately should facilitate the acceptance and usage of this technology (Davis, 1989, 1993).

Another option to influence the perception of certain characteristics of an artificial entity is to present it as acting autonomously or as being pre-programmed/pre-instructed (study IV, research paper III; study V, research paper IV). An artificial entity believed to be acting autonomously is perceived to possess more agency, competence, and responsibility for its behavior (study IV). On the other hand, enhancing that an artificial entity's behavior is determined by previous programing or instructions leads to a reduction in perceived agency (study IV; study V). However, it has to be noted that even when an artificial entity is perceived to possess low agency, i.e., low autonomy, intentionality, and responsibility (Sullins, 2006), this only affects the cognitive evaluation of certain aspects such as competence. It appears to have no substantial effect on the overall perception of the artificial entity and its behavior (study IV; study V). Thus, enhancing the perception of agency only affects certain perceptional aspects on a rather cognitive level and has not enough efficacy to alter how an artificial entity is perceived in general.

Dautenhahn (2007) indicated in her statement that behavior is the most central influencing factor and thus should be treated with great care, which goes along with the insights of many other researchers (Burgoon & Le Poire, 1993; Reeves & Nass, 1996; Rickenberg & Reeves, 2000). This also is reflected in the presented studies, which employed an actual interaction with an artificial entity (study III; study IV; study V). In detail, the results show that low interaction skills (study III) as well as harsh, offensive feedback (study IV; study V) elicit a rather negative overall evaluation of the artificial entity and the interaction with it. High interaction skills (study III) as well as benign, positive feedback (study IV), on the other hand, result in rather positive evaluations. Thus, equipping an artificial entity with sophisticated interaction skills which allow for natural and intuitive interactions with humans should evoke or foster a positive perception of the interaction as well as the artificial entity (study IV). In a similar vein should an artificial entity be designed to adhere to politeness norms, for instance, when providing feedback (study IV). However, participants of study IV also mentioned that they did not appreciate when they were praised by the robot although they knew that they gave an incorrect or no answer at all. Thus, politeness but also accuracy is requested and both aspects should be considered when designing artificial entities, particularly their behavior. Coming to a conclusion, behavioral aspects such as elaborated interaction skills, politeness, and accuracy should facilitate higher acceptance rates (Davis, 1989, 1993) and lead to more successful interactions and collaborations between humans and artificial entities (Dautenhahn, 2007).

Summing up, based on the presented results, the focus of designers should generally be laid on the behavior of artificial entities, which promises to deliver the strongest effects on people's perceptions. Since expectations appear to have an impact as well, prevalent expectations such as common replacement fears should be considered in order to avoid designing an artificial entity's appearance or behavior in a way that activates or enhances these fears. However, since expectations are overridden by interactional behavior to a great extent (cf. Burgoon & Le Poire, 1993), the design of an artificial entity's behavior should receive the greatest amount of attention, which also should be easier to accomplish than changing people's possibly deep-routed expectations. Particularly, the behavior should be designed to be easily and clearly interpretable, since ambiguity leaves room for multiple interpretations which are a lot more difficult to influence or control (Burgoon & Hale, 1988). Altogether, as the studies' results demonstrate, artificial entities should be able to communicate as natural and error-free as possible (study III) and the adherence of social norms such as politeness should be ensured by all means (study IV; study V).

4.5. Limitations and Future Research

The studies presented in this cumulus contain some limitations, which are outlined in the following. First, it needs to be noted that all studies referred to or worked with novel technologies which are not available or not able to function to the displayed extent yet. This may have caused increased effort and attention as described by the novelty effect, which often wears off over time, particularly when the technology becomes more familiar (Clark, 1983). Thus, due to the novelty effect, the perception of the technologies used in the studies may have been confounded.

Furthermore, due to the generally cross-sectional nature of the studies, causal relationships remain unclear. Thus, longitudinal studies would be a useful extension of the current results. Long-term studies would further help to understand how people's perceptions of artificial entities are influenced in the long run and/or when interacting with an artificial entity on several occasions. Previous research already showed that contact may be able to reduce negative attitudes towards artificial entities (Wullenkord et al., 2016), which would be interesting to study over a longer period of time. Likewise, moving the experiments in the field, for instance, to people's private homes or workplaces, would deliver further insights about how artificial entities are perceived outside the lab.

Regarding the sample, the majority of participants were recruited through Internet-based technologies (Facebook, eBay Classifieds, Craigslist, Instagram, and WhatsApp), which could lead to a bias regarding technological expertise. Participants were also predominantly young, highly educated, and had a similar cultural background, which could impair the generalizability of the results. Particularly, cultural imprinting may lead to different expectations regarding artificial entities (e.g., Bartneck et al., 2006; Kaplan, 2004; D. Li et al., 2010), which could ultimately also impact people's perceptions and should therefore be considered as further potential influencing factor in future research. It should be mentioned though that the last study (study V, research paper IV) contained a half American, half German sample, which showed that effects may be similar in different countries with a similar cultural background and points to the need to draw comparisons between countries with a different cultural background (Hofstede, 2001).

Most results were based on self-reports which reflects people's cognitive reactions on a more conscious level and may be constrained or biased, for instance, with regard to social desirability (Grimm, 2010). For this reason, study IV (research paper III) included a test measuring the automatic activation of attitudes to extend self-reported measures by a behavioral component, however, no conclusive information could be obtained here. More research employing further behavioral measurements would extend the findings of the studies included in the cumulus. Nevertheless, I would like to emphasize that in three of the five presented studies a realistic interaction with a robot or virtual agent took place, which extends previous insights of survey-only studies and enabled the examination of the perception of artificial entities in realistic settings.

With regards to potential influencing factors for the perception of artificial entities, the studies presented here examine a selection of factors which were assumed to have substantial effects. However, there are several other factors, on side of the user (e.g., gender, social influences, and socioeconomic status; Grint & Gill, 1995; van Dijk, 2006; Venkatesh & Davis, 2000) as well as on side of the artificial entity (e.g., appearance, voice, and facial expressions; Eyssel et al., 2012; Rosenthal-von der Pütten et al., 2018; Syrdal et al., 2007) which should be explored in more depth in future studies. Particularly since the artificial entity's behavior proved to be the strongest influencing factor, further behavioral variations and their effects should be studied. Of particular interest would be to examine whether expectation and attribution processes play a stronger role when an artificial entity's behavior is ambiguous and thus allows for several interpretations. Moreover, influencing factors could also be considered separately for social robots and virtual agents in future studies since there are some differences, for instance, regarding the physical expressiveness (Holz et al., 2009). However, both forms of artificial entity were found to generally evoke similar human reactions in previous research (Fridin & Belokopytov, 2014; Holz et al., 2009; J. Li et al., 2016) and these tendencies could also be observed when comparing study IV and study V.

4.6. Conclusion

Against the background of the rapid development of artificial entities and their various effects on their human interaction partners, the overarching aim of the studies presented in this cumulus was to examine different factors with regard to their influence on the perception of these artificial entities. Overall, some factors were identified that influence the perception of certain aspects, while one factor was found to have an overall effect on the perception. In general, expectations appear to form a perceptual frame, which is further influenced by people's technological expertise. Expectations reflecting the two most popular and simultaneously conflicting views on artificial entities – that of a helpful assistant versus an intimidating competitor – influence how sociable people perceive an artificial entity. Furthermore, the extent to which an artificial entity is portrayed as acting autonomously influences its perceived agency, responsibility, and competence as well as the degree to which it elicits affectional responses.

However, the factor with the greatest effect sizes influencing all kinds of evaluation measures was the artificial entity's behavior. Particularly, behavior which can be clearly classified as desirable or repulsive has a great impact on perceiving the artificial entity in a positive or negative light. These findings confirm and extend previous research where the actual behavior of a person (Burgoon & Le Poire, 1993), a computer (Nass & Moon, 2000), or a virtual agent (Rickenberg & Reeves, 2000) was shown to have a substantial effect on people's perception, by demonstrating its relevance within a framework of other influencing factors such as expectations, technological expertise, and attribution processes. Altogether, an artificial entity's actual behavior was found to exceed, sometimes even overshadow, those other influences and also appeared to be unaffected by them.

In order to successfully design pleasant artificial entities and desirable interactions with them, factors influencing the perception of artificial entities should be kept in mind. Particularly, technological expertise and positive expectations should be fostered while avoiding to evoke or enhance negative expectations or even fears. To enhance the perception of agency and competence, the autonomy of the artificial entity should be emphasized. Eventually, most consideration should be spent on the careful construction and design of an artificial entity's behavior depending on which perception should be elicited, whereby impolite behavior as well as poor interaction skills should be generally avoided.

Furthermore, this work helps to expand our general knowledge regarding the human perception of an interaction partner. Particularly by examining how artificial entities are perceived we receive a deepened understanding of the human mind since it teaches us which specific factors are decisive for what kinds of perceptions. By examining those factor variations within interactions with machines, we are able to consider and subsequently single out their specific influence while systematically adding or removing the influence of other factors. Thus, altogether we learned that expectations and attributions affect the perception of certain characteristics of an interaction partner on a rather conscious, cognitive level, while behavior with a clear, unambiguous meaning has the greatest effect on the overall perception, without being affected by assumptions based on expectations or attributions.

5. REFERENCES

Allen, C., Wallach, W., & Smit, I. (2006). Why machine ethics? *IEEE Intelligent Systems*, *21*(4), 12–17. https://doi.org/10.1109/MIS.2006.83

Andersen, S. M., & Klatzky, R. L. (1987). Traits and social stereotypes: Levels of categorization in person perception. *Journal of Personality and Social Psychology*, *53*(2), 235–246. https://doi.org/10.1037/0022-3514.53.2.235

Arras, K. O., & Cerqui, D. (2005). Do we want to share our lives and bodies with robots? A 2000 people survey. *Technical Report Nr. 0605-001, Autonomous Systems Lab, Swiss Federal Institute of Technology, EPFL*, *2005.* https://doi.org/10.3929/ETHZ-A-010113633

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Asher, D. E., Zaldivar, A., Barton, B., Brewer, A. A., & Krichmar, J. L. (2012). Reciprocity and retaliation in social games with adaptive agents. *IEEE Transactions on Autonomous Mental Development*, *4*(3), 226–238. https://doi.org/10.1109/TAMD.2012.2202658

Asimov, I. (1947). *Little lost robot*. Street & Smith.

Asimov, I. (1950). *I, robot*. Gnome Press.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, *94*(2), 416–427. https://doi.org/10.1037/0022-0663.94.2.416

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, *31*(2), 101–119. https://doi.org/10.1080/0305724022014322

Banks, J. (2019). A perceived moral agency scale: Development and validation of a metric for humans and social machines. *Computers in Human Behavior*, *90*, 363–371. https://doi.org/10.1016/j.chb.2018.08.028

Banks, J. (2020). Optimus primed: Media cultivation of robot mental models and social judgments. *Frontiers in Robotics and AI*, *7*, 62. https://doi.org/10.3389/frobt.2020.00062

Bannister, B. D. (1986). Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. *Journal of Applied Psychology*, *71*(2), 203–210. https://doi.org/10.1037/0021-9010.71.2.203

Barker, C. (2005). *Cultural studies: Theory and practice*. Sage.

Bartneck, C. (2004). From fiction to science – A cultural reflection of social robots. In J. Forlizzi & C. Bartneck (Eds.), *Proceedings of the CHI 2004 Workshop on Shaping Human-Robot Interaction: Understanding the Social Aspects of Intelligent Robotic Products* (pp. 1–4). ACM Press.

Bartneck, C., & Forlizzi, J. (2004). A design-centred framework for social human-robot interaction. In *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN '04* (pp. 591–594). IEEE. https://doi.org/10.1109/ROMAN.2004.1374827

Bartneck, C., Kanda, T., Mubin, O., & Mahmud, A. A. (2007). The perception of animacy and intelligence based on a robot's embodiment. In J. Kuffner (Ed.), *Proceedings of the 7th IEEE-RAS International Conference on Humanoid Robots - Humanoids '07* (pp. 300–305). IEEE. https://doi.org/10.1109/ICHR.2007.4813884

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Bartneck, C., Suzuki, T., Kanda, T., & Nomura, T. (2006). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & Society*, *21*(1-2), 217–230. https://doi.org/10.1007/s00146-006-0052-7

Bartneck, C., van der Hoek, M., Mubin, O., & Al Mahmud, A. (2007). "Daisy, Daisy, give me your answer do!": Switching off a robot. In C. Breazeal (Ed.), *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction - HRI '07* (pp. 217–222). ACM. https://doi.org/10.1145/1228716.1228746

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Baylor, A. L., Ryu, J., & Shen, E. (2003). The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. In D. Lassner & C. McNaught (Eds.), *Proceedings of the ED-MEDIA 2003 World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 452–458). AACE.

Beier, G. (1999). Kontrollüberzeugungen im Umgang mit Technik [Locus of control when using technology]. *Report Psychologie*, *9*, 684–693.

Berge, Z. L., Muilenburg, L. Y., & van Haneghan, J. (2002). Barriers to distance education and training: Survey results. *Quarterly Review of Distance Education*, *3*(4), 409–419.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, *1*(2), 99–112. https://doi.org/10.1111/j.1468-2958.1975.tb00258.x

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *Intelligent Virtual Agents: Proceedings of the 12th International Conference on Intelligent Virtual Agents - IVA '12* (Vol. 7502, pp. 126–138). Springer. https://doi.org/10.1007/978-3-642-33197-8_13

Berlo, D. K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for evaluating the acceptability of message sources. *Public Opinion Quarterly*, *33*(4), 563–576. https://doi.org/10.1086/267745

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Bittermann, A., Kühnlenz, K., & Buss, M. (2007). On the evaluation of emotion expressing robots. In *Proceedings of the 2007 IEEE International Conference on Robotics and Automation - ICRA '07* (pp. 2138–2143). IEEE. https://doi.org/10.1109/ROBOT.2007.363637

Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. In W. Broll, C. Greenhalgh, & E. F. Churchill (Eds.), *Proceedings of the 4th International Conference on Collaborative Virtual Environments - CVE '02* (pp. 25–30). ACM Press. https://doi.org/10.1145/571878.571883

Blascovich, J., Loomis, J. M., 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*(2), 103–124. https://doi.org/10.1207/S15327965PLI1302_01

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, *63*(2), 131–144. https://doi.org/10.1006/obhd.1995.1068

Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: Categorizing social interactions. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Evolution and cognition: Simple heuristics that make us smart* (pp. 257–285). Oxford University Press.

Breazeal, C. (2002). *Designing sociable robots*. MIT press.

Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, *42*(3-4), 167–175. https://doi.org/10.1016/S0921-8890(02)00373-1

Breazeal, C. (2011). Social robots for health applications. In P. Bonato, C. Brenan, A. Laine, & M. Akay (Eds.), *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 2011, pp. 5368–5371). IEEE. https://doi.org/10.1109/IEMBS.2011.6091328

Breazeal, C., & Scassellati, B. (1999). How to build robots that make friends and influence people. In *Proceedings of the12th IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS '99* (pp. 858–863). IEEE. https://doi.org/10.1109/IROS.1999.812787

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: a review. *Gerontechnology*, *8*(2), 94–103. https://doi.org/10.4017/gt.2009.08.02.002.00

Brown, L. A., & Walker, W. H. (2008). Prologue: Archaeology, animism and non-human agents. *Journal of Archaeological Method and Theory*, *15*(4), 297–299. https://doi.org/10.1007/s10816-008-9056-6

Bruce, A., Nourbakhsh, I., & Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation* (pp. 4138–4142). IEEE. https://doi.org/10.1109/ROBOT.2002.1014396

Bruckenberger, U., Weiss, A., Mirnig, N., Strasser, E., Stadler, S., & Tscheligi, M. (2013). The good, the bad, the weird: Audience evaluation of a “real” robot in relation to science fiction and mass media. In M. Jamshidi (Ed.), *Advance Trends in Soft Computing: Proceedings of the World Conference on Soft Computing - WCSC '13* (Vol. 8239, pp. 301–310). Springer. https://doi.org/10.1007/978-3-319-02675-6_30

Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., & Thrun, S. (1999). Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, *114*(1-2), 3–55. https://doi.org/10.1016/S0004-3702(99)00070-3

Burgoon, J. K. (1993). Interpersonal Expectations, Expectancy Violations, and Emotional Communication. *Journal of Language and Social Psychology*, *12*(1-2), 30–48. https://doi.org/10.1177/0261927X93121003

Burgoon, J. K., Bonito, J. A., Lowry, P. B., Humpherys, S. L., Moody, G. D., Gaskin, J. E., & Giboney, J. S. (2016). Application of Expectancy Violations Theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human-Computer Studies*, *91*, 24–36. https://doi.org/10.1016/j.ijhcs.2016.02.002

Burgoon, J. K., & Hale, J. L. (1988). Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors. *Communication Monographs*, *55*(1), 58–79. https://doi.org/10.1080/03637758809376158

Burgoon, J. K., & Jones, S. B. (1976). Toward a theory of personal space expectations and their violations. *Human Communication Research*, *2*(2), 131–146. https://doi.org/10.1111/j.1468-2958.1976.tb00706.x

Burgoon, J. K., & Le Poire, B. A. (1993). Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. *Human Communication Research*, *20*(1), 67–96. https://doi.org/10.1111/j.1468-2958.1993.tb00316.x

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research*, *17*(2), 232–265. https://doi.org/10.1111/j.1468-2958.1990.tb00232.x

Carolus, A., Muench, R., Schmidt, C., & Schneider, F. (2019). Impertinent mobiles - Effects of politeness and impoliteness in human-smartphone interaction. *Computers in Human Behavior*, *93*, 290–300. https://doi.org/10.1016/j.chb.2018.12.030

Cassell, J., & Thórisson, K. R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, *13*(4-5), 519–538. https://doi.org/10.1080/088395199117360

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1226. https://doi.org/10.1126/science.1165565

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*(4), 445–459. https://doi.org/10.3102/00346543053004445

Clarke, A. C. (1968). *2001: A space odyssey*. New American Library.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies: Why and how. *Knowledge-Based Systems*, *6*(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist*, *35*(10), 867–881. https://doi.org/10.1037/0003-066X.35.10.867

Darling, K., Nandy, P., & Breazeal, C. (2015). Empathic concern and the effect of stories in human-robot interaction. In *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '15* (pp. 770–775). IEEE. https://doi.org/10.1109/ROMAN.2015.7333675

Dautenhahn, K. (1999). Embodiment and interaction in socially intelligent life-like agents. In C. L. Nehaniv (Ed.), *Proceedings of the International Workshop on Computation for Metaphors, Analogy, and Agents - CMAA '98* (pp. 102–141). Springer. https://doi.org/10.1007/3-540-48834-0_7

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society of London*, *362*(1480), 679–704. https://doi.org/10.1098/rstb.2006.2004

Dautenhahn, K., Woods, S. N., Kaouri, C., Walters, M. L., Koay, K. L., & Werry, I. (2005). What is a robot companion - Friend, assistant or butler? In *Proceedings of the 18th*

IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS '05 (pp. 1192–1197). IEEE. https://doi.org/10.1109/IROS.2005.1545189

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–339. https://doi.org/10.2307/249008

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, *38*(3), 475–487. https://doi.org/10.1006/imms.1993.1022

de Charms, R. (1983). *Personal causation: The internal affective determinants of behavior*. Lawrence Erlbaum Associates, Inc.

de Melo, C. M., Marsella, S., & Gratch, J. (2016). "Do as I say, not as I do" Challenges in delegating decisions to automated agents. In J. Thangarajah, K. Tuyls, C. M. Jonker, & S. Marsella (Eds.), *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems - AAMAS '16* (pp. 949–956). IFAAMAS.

de Melo, C. M., Marsella, S., & Gratch, J. (2018). Social decisions and fairness change when people's interests are represented by autonomous agents. *Autonomous Agents and Multi-Agent Systems*, *32*(1), 163–187. https://doi.org/10.1007/s10458-017-9376-6

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *53*(6), 1024–1037. https://doi.org/10.1037/0022-3514.53.6.1024

Dick, P. K. (1968). *Do androids dream of electric sheep?* Doubleday.

Embgen, S., Luber, M., Becker-Asano, C., Ragni, M., Evers, V., & Arras, K. O. (2012). Robot-specific social cues in emotional body language. In *Proceedings of the 21st IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN '12* (pp. 1019–1025). IEEE. https://doi.org/10.1109/ROMAN.2012.6343883

Emirbayer, M., & Mische, A. (1998). What Is Agency? *American Journal of Sociology*, *103*(4), 962–1023. https://doi.org/10.1086/231294

Eyssel, F., de Ruiter, L., Kuchenbrandt, D., Bobinger, S., & Hegel, F. (2012). 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In H. Yanco, A. Steinfeld, V. Evers, & O. C. Jenkins (Eds.), *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction - HRI '12* (pp. 125–126). IEEE. https://doi.org/10.1145/2157689.2157717

Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, *42*(9), 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

Ezer, N., Fisk, A. D., & Rogers, W. A. (2009). Attitudinal and intentional acceptance of domestic robots by younger and older adults. In C. Stephanidis (Ed.), *Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction - UAHCI '09* (Vol. 5615, pp. 39–48). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02710-9_5

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis.

Fogg, B. J., & Nass, C. I. (1997a). How users reciprocate to computers: An experiment that demonstrates behavior change. In A. Edwards & S. Pemberton (Eds.), *CHI '97 extended* *abstracts on Human factors in computing systems looking to the future - CHI '97* (pp. 331–332). ACM Press. https://doi.org/10.1145/1120212.1120419

Fogg, B. J., & Nass, C. I. (1997b). Silicon sycophants: the effects of computers that flatter. *International Journal of Human-Computer Studies*, *46*(5), 551–561. https://doi.org/10.1006/ijhc.1996.0104

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

Fox, J., Ahn, S. J., Janssen, J. H., Yeykelis, L., Segovia, K. Y., & Bailenson, J. N. (2015). Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence. *Human–Computer Interaction*, *30*(5), 401–432. https://doi.org/10.1080/07370024.2014.921494

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human–Computer Interaction*, *35*(6), 456–467. https://doi.org/10.1080/10447318.2018.1456150

Fridin, M., & Belokopytov, M. (2014). Embodied robot versus virtual agent: Involvement of preschool children in motor task performance. *International Journal of Human–Computer Interaction*, *30*(6), 459–469. https://doi.org/10.1080/10447318.2014.888500

Gaul, S., Ziefle, M., Arning, K., Wilkowska, W., Kasugai, K., Röcker, C., & Jakobs, E.-M. (2010). Technology acceptance as an integrative component of product developments in the medical technology sector. In *Proceedings of the 3rd Ambient Assisted Living Conference - AAL '10* (pp. 26–27). VDE-Verlag.

Gerbner, G., & Gross, L. (1976). Living with television: The violence profile. *Journal of Communication*, *26*(2), 173–199. https://doi.org/10.1111/j.1460-2466.1976.tb01397.x

Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A. C., & Wang, J. (2005). Designing robots for long-term social interaction. In *Proceedings of the 18th IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS '05* (pp. 1338–1343). IEEE. https://doi.org/10.1109/IROS.2005.1545303

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN '03* (pp. 55–60). IEEE. https://doi.org/10.1109/ROMAN.2003.1251796

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.

Gogoll, J., & Uhl, M. (2018). Rage against the machine: Automation in the moral domain. *Journal of Behavioral and Experimental Economics*, *74*, 97–103. https://doi.org/10.1016/j.socec.2018.04.003

Gratch, J., DeVault, D., & Lucas, G. M. (2016). The benefits of virtual humans for teaching negotiation. In D. Traum, W. Swartout, P. Khooshabeh, S. Kopp, S. Scherer, & A. Leuski (Eds.), *Intelligent Virtual Agents: Proceedings of the 16th International Conference on Intelligent Virtual Agents - IVA '16* (pp. 283–294). Springer. https://doi.org/10.1007/978-3-319-47665-0_25

Grimm, P. (2010). Social desirability bias. In J. N. Sheth & N. K. Malhotra (Eds.), *Wiley international encyclopedia of marketing.* John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444316568.wiem02057

Grint, K., & Gill, R. (Eds.). (1995). *The gender-technology relation: Contemporary theory and research*. Taylor & Francis.

Gunkel, D. J. (2014). A vindication of the rights of machines. *Philosophy & Technology*, *27*(1), 113–132. https://doi.org/10.1007/s13347-013-0121-z

Hancock, P. A. (2017). Imposing limits on autonomous systems. *Ergonomics*, *60*(2), 284–291. https://doi.org/10.1080/00140139.2016.1190035

Häring, M., Bee, N., & André, E. (2011). Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '11* (pp. 204–209). IEEE. https://doi.org/10.1109/ROMAN.2011.6005263

Heerink, M., Kröse, B. J. A., Wielinga, B., & Evers, V. (2009). Measuring the influence of social abilities on acceptance of an interface robot and a screen agent by elderly users. In A. F. Blackwell (Ed.), *Electronic Workshops in Computing, People and Computers XXIII Celebrating People and Technology: Proceedings of the 11th International Conference on Human-Computer Interaction - HCI '09* (pp. 751–756). BCS Learning & Development. https://doi.org/10.14236/ewic/HCI2009.54

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243–259. https://doi.org/10.2307/1416950

Hewstone, M., & Jaspars, J. (1987). Covariation and causal attribution: A Logical Model of the intuitive analysis of variance. *Journal of Personality and Social Psychology*, *53*(4), 663–672. https://doi.org/10.1037/0022-3514.53.4.663

Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent? *Ethics and Information Technology*, *11*(1), 19–29. https://doi.org/10.1007/s10676-008-9167-5

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, *19*(1-2), 151–181. https://doi.org/10.1080/07370024.2004.9667343

Hoffmann, L., Krämer, N. C., Lam-chi, A., & Kopp, S. (2009). Media equation revisited: Do users show polite reactions towards an embodied agent? In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Eds.), *Intelligent Virtual Agents: Proceedings of the 9th International Conference on Intelligent Virtual Agents - IVA '09* (pp. 159–165). Springer. https://doi.org/10.1007/978-3-642-04380-2_19

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage Publications, Inc.

Holz, T., Dragone, M., & O'Hare, G. M. P. (2009). Where robots and virtual agents meet: A survey of social interaction research across Milgram's reality-virtuality continuum. *International Journal of Social Robotics*, *1*(1), 83–93. https://doi.org/10.1007/s12369-008-0002-2

Honeycutt, J. M. (1991). Role of nonverbal behaviors in modifying expectancies during initial encounters. *Southern Communication Journal*, *56*(3), 161–177. https://doi.org/10.1080/10417949109372828

Horstmann, A. C., Bock, N., Linhuber, E., Szczuka, J. M., Straßmann, C., & Krämer, N. C. (2018). Do a robot's social skills and its objection discourage interactants from switching the robot off? *PloS One*, *13*(7), e0201581. https://doi.org/10.1371/journal.pone.0201581

Horstmann, A. C., & Krämer, N. C. (2019). Great expectations? Relation of previous experiences with social robots in real life or in the media and expectancies based on qualitative and quantitative assessment. *Frontiers in Psychology*, *10*, 939. https://doi.org/10.3389/fpsyg.2019.00939

Horstmann, A. C., & Krämer, N. C. (2020). Expectations vs. Actual behavior of a social robot: An experimental investigation of the effects of a social robot's interaction skill level and its expected future role on people's evaluations. *PloS One*, *15*(8), e0238133. https://doi.org/10.1371/journal.pone.0238133

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.

Huber, M., Rickert, M., Knoll, A., Brandt, T., & Glasauer, S. (2008). Human-robot interaction in handing-over tasks. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '08* (pp. 107–112). IEEE. https://doi.org/10.1109/ROMAN.2008.4600651

Ickes, W., Patterson, M. L., Rajecki, D. W., & Tanford, S. (1982). Behavioral and cognitive consequences of reciprocal versus compensatory responses to preinteraction expectancies. *Social Cognition*, *1*(2), 160–190. https://doi.org/10.1521/soco.1982.1.2.160

Isbister, K., & Nass, C. I. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, *53*(2), 251–267. https://doi.org/10.1006/ijhc.2000.0368

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., & Nakatsu, R. (2001). Robovie: An interactive humanoid robot. *Industrial Robot: An International Journal*, *28*(6), 498–504. https://doi.org/10.1108/01439910110410051

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 219–266). Academic Press. https://doi.org/10.1016/S0065-2601(08)60107-0

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*(1), 1–24. https://doi.org/10.1016/0022-1031(67)90034-0

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–95). General Learning Press.

Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J. H., & Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. In A. Billard, P. H. Kahn, J. A. Adams, & G. Trafton (Eds.), *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction - HRI '11* (pp. 159–160). ACM Press. https://doi.org/10.1145/1957656.1957710

Kang, S.-H., Gratch, J., Wang, N., & Watt, J. H. (2008). Agreeable people like agreeable virtual humans. In H. Prendinger, J. C. Lester, & M. Ishizuka (Eds.), *Intelligent Virtual Agents: Proceedings of the 8th International Conference on Intelligent Virtual Agents - IVA '08* (pp. 253–261). Springer. https://doi.org/10.1007/978-3-540-85483-8_26

Kaplan, F. (2004). Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International Journal of Humanoid Robotics*, *01*(03), 465–480. https://doi.org/10.1142/S0219843604000289

Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinität erfassen: Der Fragebogen TA-EG [Measuring technical affinity - the questionnaire TA-EG]. *Der Mensch Im Mittelpunkt Technischer Systeme*, *8*, 196–201.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107–128. https://doi.org/10.1037/h0034225

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457–501. https://doi.org/10.1146/annurev.ps.31.020180.002325

Khan, Z. (1998). *Attitudes towards intelligent service robots*. Royal Institute of Technology.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049. https://doi.org/10.1007/s10803-012-1645-2

Kim, T., & Hinds, P. J. (2006). Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '06* (pp. 80–85). IEEE. https://doi.org/10.1109/ROMAN.2006.314398

Kiss, G. (1992). Variable coupling of agents to their environment: Combining situated and symbolic automata. In E. Werner & Y. Demazeau (Eds.), *Decentralized Artificial Intelligence: Proceedings of the Modelling of Autonomous Agents in Multi-Agent Worlds Conference - MAAMAW '91* (pp. 231–237). Elsevier.

Komatsu, T., & Yamada, S. (2011). Adaptation gap hypothesis: How differences between users' expected and perceived agent functions affect their subjective impression. *Journal of Systemics, Cybernetics and Informatics*, *9*(1), 67–74.

Kozima, H., Nakagawa, C., & Yano, H. (2003). Attention coupling as a prerequisite for social interaction. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN '03* (pp. 109–114). IEEE. https://doi.org/10.1109/ROMAN.2003.1251814

Krämer, N. C., Leiße, L.-M., Hollingshead, A., & Gratch, J. (2017). Evaluated by a machine: Effects of negative feedback by a computer or human boss. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Intelligent Virtual Agents: Proceedings of the 17th International Conference on Intelligent Virtual Agents - IVA '17* (pp. 235–238). Springer. https://doi.org/10.1007/978-3-319-67401-8_29

Krämer, N. C., von der Pütten, A. M., & Eimler, S. C. (2012). Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In M. Zacarias & J. V. de Oliveira (Eds.), *Studies in Computational Intelligence. Human-computer interaction: The agency perspective* (Vol. 396, pp. 215–240). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-25691-2_9

Kriz, S., Ferro, T. D., Damera, P., & Porter, J. R. (2010). Fictional robots as a data source in HRI research: Exploring the link between science fiction and interactional expectations. In *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '10* (pp. 458–463). IEEE. https://doi.org/10.1109/ROMAN.2010.5598620

Kühnlenz, K., Sosnowski, S., & Buss, M. (2010). Impact of animal-like features on emotion expression of robot head EDDIE. *Advanced Robotics*, *24*(8-9), 1239–1255. https://doi.org/10.1163/016918610X501309

Kwon, M., Jung, M. F., & Knepper, R. A. (2016). Human expectations of social robots. In C. Bartneck, Y. Nagai, A. Paiva, & S. Sabanovic (Eds.), *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction - HRI '16* (pp. 463–464). IEEE. https://doi.org/10.1109/HRI.2016.7451807

Lane, C. A., & Lyle, H. F. (2011). Obstacles and supports related to the use of educational technologies: The role of technological expertise, gender, and age. *Journal of Computing in Higher Education*, *23*(1), 38–59. https://doi.org/10.1007/s12528-010-9034-3

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication*, *56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Lee, S., Lau, I. Y., Kiesler, S., & Chiu, C.-Y. (2005). Human mental models of humanoid robots. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation - ICRA '05* (pp. 2767–2772). IEEE. https://doi.org/10.1109/ROBOT.2005.1570532

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: A survey. *International Journal of Social Robotics*, *5*(2), 291–308. https://doi.org/10.1007/s12369-013-0178-y

Lester, J. C., Towns, S. G., Callaway, C. B., Voerman, J. L., & FitzGerald, P. J. (2000). Deictic and emotive communication in animated pedagogical agents. In J. Cassell, J. Sullivan, S. Prevost, & E. F. Churchill (Eds.), *Embodied conversational agents.* The MIT Press. https://doi.org/10.7551/mitpress/2697.003.0007

Leviathan, Y., & Matias, Y. (2018). *Google Duplex: An AI system for accomplishing real-world tasks over the phone.* Google AI Blog. https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html

Li, D., Rau, P. L. P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics*, *2*(2), 175–186. https://doi.org/10.1007/s12369-010-0056-9

Li, J., Kizilcec, R., Bailenson, J. N., & Ju, W. (2016). Social robots and virtual agents as lecturers for video instruction. *Computers in Human Behavior*, *55*, 1222–1230. https://doi.org/10.1016/j.chb.2015.04.005

Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, *123*(1), 3–32. https://doi.org/10.1037/0033-2909.123.1.3

Looije, R., Neerincx, M. A., & Cnossen, F. (2010). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies*, *68*(6), 386–397. https://doi.org/10.1016/j.ijhcs.2009.08.007

Lorenz, T., Weiss, A., & Hirche, S. (2016). Synchrony and reciprocity: Key mechanisms for social companion robots in therapy and care. *International Journal of Social Robotics*, *8*(1), 125–143. https://doi.org/10.1007/s12369-015-0325-8

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies*, *16*(2), 141–172. https://doi.org/10.1075/is.16.2.01mac

Malle, B. F., & Hodges, S. D. (Eds.). (2007). *Other minds: How humans bridge the divide between self and others*. Guilford Press.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*(2), 101–121. https://doi.org/10.1006/jesp.1996.1314

Malle, B. F., Moses, L. J., & Baldwin, D. A. (Eds.). (2001). *Intentions and intentionality: Foundations of social cognition*. MIT press.

McCauley, L. (2007). The Frankenstein Complex and Asimov's three laws. In T. Metzler (Ed.), *Papers from the 2007 AAAI Workshop* (pp. 9–14). AAAI Press.

Mirnig, N., Strasser, E., Weiss, A., & Tscheligi, M. (2012). Studies in public places as a means to positively influence people's attitude towards robots. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, S. S. Ge, O. Khatib, J.-J. Cabibihan, R. Simmons, & M.-A. Williams (Eds.), *Social robotics* (pp. 209–218). Springer. https://doi.org/10.1007/978-3-642-34103-8_21

Morahan-Martin, J., & Schumacher, P. (2007). Attitudinal and experiential predictors of technological expertise. *Computers in Human Behavior*, *23*(5), 2230–2239. https://doi.org/10.1016/j.chb.2006.03.003

Mori, M. (1970). The uncanny valley. *Energy*, *7*(4), 33–35.

Morkes, J., Kernal, H. K., & Nass, C. I. (1999). Effects of humor in task-oriented human-computer interaction and computer-mediated communication: A direct test of SRCT theory. *Human–Computer Interaction*, *14*(4), 395–435. https://doi.org/10.1207/S15327051hci1404_2

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1997). Automation bias: Decision making and performance in high-tech cockpits. *The International Journal of Aviation Psychology*, *8*(1), 47–63. https://doi.org/10.1207/s15327108ijap0801_3

Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In A. Billard, P. H. Kahn, J. A. Adams, & G. Trafton (Eds.), *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction - HRI '11* (pp. 331–338). ACM Press. https://doi.org/10.1145/1957656.1957786

Murphy, R. R. (2018). Astromech robots in Star Wars. *Science Robotics*, *3*(15), eaat1599. https://doi.org/10.1126/scirobotics.aat1599

Nass, C. I., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*(6), 669–678. https://doi.org/10.1006/ijhc.1996.0073

Nass, C. I., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, *7*(3), 171–181. https://doi.org/10.1037/1076-898X.7.3.171

Nass, C. I., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C. I., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, *29*(5), 1093–1109. https://doi.org/10.1111/j.1559-1816.1999.tb00142.x

Nass, C. I., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, *27*(10), 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Nisbet, M. C., Scheufele, D. A., Shanahan, J., Moy, P., Brossard, D., & Lewenstein, B. V. (2002). Knowledge, reservations, or promise? A media effects model for public perceptions of science and technology. *Communication Research*, *29*(5), 584–608. https://doi.org/10.1177/009365002236196

Nisbett, R. E., & Ross, L. (1983). *Human inference: Strategies and shortcomings of social judgment*. Prentice Hall.

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human-robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, *24*(2), 442–451. https://doi.org/10.1109/TRO.2007.914004

Nomura, T., Sugimoto, K., Syrdal, D. S., & Dautenhahn, K. (2012). Social acceptance of humanoid robots in Japan: A survey for development of the Frankenstein Syndrome Questionnaire. In *Proceedings of the 12th IEEE-RAS International Conference on Humanoid Robots* (pp. 242–247). IEEE. https://doi.org/10.1109/HUMANOIDS.2012.6651527

Nuñez, T. R., & Rosenthal-von der Pütten, A. M. (2018). Roboter und Agenten in der psychologischen Intervention [Robots and agents in the psychological intervention]. In O. D. Kothgassner & A. Felnhofer (Eds.), *Klinische Cyberpsychologie und Cybertherapie* (pp. 78–85). Facultas Universitätsverlag.

Ochs, M., Niewiadomski, R., & Pelachaud, C. (2010). How a virtual agent should smile? Morphological and dynamic characteristics of virtual agent's smiles. In J. Allbeck, . Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Intelligent Virtual Agents: Proceedings of the 10th International Conference on Intelligent Virtual Agents - IVA '10* (pp. 427–440). Springer. https://doi.org/10.1007/978-3-642-15892-6_47

Oestreicher, L., & Eklundh, K. (2006). User expectations on human-robot co-operation. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '06* (pp. 91–96). IEEE. https://doi.org/10.1109/ROMAN.2006.314400

Oracle. (2016). *Can virtual experiences replace reality? The future role for humans in delivering customer experience*. Redwood Shores, CA.

Paiva, A., & Machado, I. (1998). Vincent, an autonomous pedagogical agent for on-the-job training. In B. P. Goettl, H. M. Halff, C. L. Redfield, & V. J. Shute (Eds.), *Proceedings of the 4th International Conference on Intelligent Tutoring Systems - ITS '98* (pp. 584–593). Springer. https://doi.org/10.1007/3-540-68716-5_64

Parise, S., Kiesler, S., Sproull, L., & Waters, K. (1999). Cooperating with life-like interface agents. *Computers in Human Behavior*, *15*(2), 123–142. https://doi.org/10.1016/S0747-5632(98)00035-1

Parisot, A. H. (1997). Distance education as a catalyst for changing teaching in the community college: Implications for institutional policy. *New Directions for Community Colleges*, *99*, 5–13. https://doi.org/10.1002/cc.9901

Pelachaud, C. (2009a). Modelling multimodal expression of emotion in a virtual agent. *Philosophical Transactions of the Royal Society of London*, *364*(1535), 3539–3548. https://doi.org/10.1098/rstb.2009.0186

Pelachaud, C. (2009b). Studies on gesture expressivity for a virtual agent. *Speech Communication*, *51*(7), 630–639. https://doi.org/10.1016/j.specom.2008.04.009

PwC. (2018). *Consumer Intelligence Series: Prepare for the voice revolution.*

Ramey, C. H. (2006). An inventory of reported characteristics for home computers, robots, and human beings: Applications for android science and the uncanny valley. In *Proceedings of the 5th International Conference of the Cognitive Science - ICCS '06* (pp. 43–47). Cognitive Science Society.

Ray, C., Mondada, F., & Siegwart, R. (2008). What do people expect from robots? In *Proceedings of the 21st IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS '08* (pp. 3816–3821). IEEE. https://doi.org/10.1109/IROS.2008.4650714

Reeder, G. D. (2009). Mindreading: Judgments about intentionality and motives in dispositional inference. *Psychological Inquiry*, *20*(1), 1–18. https://doi.org/10.1080/10478400802615744

Reeder, G. D., Kumar, S., Hesson-McInnis, M., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, *83*(4), 789–803. https://doi.org/10.1037//0022-3514.83.4.789

Reeves, B., Hancock, J., & Liu, X. (2020). Social robots are like real people: First impressions, attributes, and stereotyping of social robots. *Technology, Mind, and Behavior*, *1*(1), 1–14. https://doi.org/10.1037/tmb0000018

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. CSLI Publications.

Rickenberg, R., & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In T. Turner (Ed.), *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 49–56). ACM. https://doi.org/10.1145/332040.332406

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., & Robinson, P. (2009). How anthropomorphism affects empathy toward robots. In M. Scheutz, F. Michaud, P. J. Hinds, & B. Scassellati (Eds.), *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction - HRI '09* (pp. 245–246). IEEE. https://doi.org/10.1145/1514095.1514158

Rose, J., & Truex, D. (2000). Machine agency as perceived autonomy: An action perspective. In R. Baskerville, J. Stage, & J. I. DeGross (Eds.), *Organizational and social perspectives on information technology* (pp. 371–388). Springer. https://doi.org/10.1007/978-0-387-35505-4_22

Rosenthal-von der Pütten, A. M., & Krämer, N. C. (2015). Individuals' evaluations of and attitudes towards potentially uncanny robots. *International Journal of Social Robotics*, *7*(5), 799–824. https://doi.org/10.1007/s12369-015-0321-z

Rosenthal-von der Pütten, A. M., Krämer, N. C., & Herrmann, J. (2018). The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics*, *10*(5), 569–582. https://doi.org/10.1007/s12369-018-0466-7

Rosenthal-von der Pütten, A. M., Straßmann, C., & Krämer, N. C. (2016). Robots or agents – Neither helps you more or less during second language acquisition. In D. Traum, W. Swartout, P. Khooshabeh, S. Kopp, S. Scherer, & A. Leuski (Eds.), *Intelligent Virtual Agents: Proceedings of the 16th International Conference on Intelligent Virtual Agents - IVA '16* (pp. 256–268). Springer. https://doi.org/10.1007/978-3-319-47665-0_23

Rosenthal-von der Pütten, A. M., & Weiss, A. (2015). The uncanny valley phenomenon: Does it affect all of us? *Interaction Studies*, *16*(2), 206–214. https://doi.org/10.1075/is.16.2.07ros

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 173–220. https://doi.org/10.1016/S0065-2601(08)60357-3

Sandoval, E. B., Mubin, O., & Obaid, M. (2014). Human robot interaction and fiction: A contradiction. In M. Beetz (Ed.), *Proceedings of the 6th International Conference on Social Robotics - ICSR '14* (Vol. 8755, pp. 54–63). Springer. https://doi.org/10.1007/978-3-319-11973-1_6

Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H., & Hagita, N. (2009). How to approach humans? Strategies for social robots to initiate interaction. In M. Scheutz, F. Michaud, P. J. Hinds, & B. Scassellati (Eds.), *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction - HRI '09* (pp. 109–116). IEEE. https://doi.org/10.1145/1514095.1514117

Sayin, E., & Krishna, A. (2019). You can't be too polite, Alexa! Implied politeness of mechanized auditory feedback and its impact on perceived performance accuracy. In R. Bagchi, L. Block, & L. Lee (Eds.), *Advances in consumer research* (Vol. 47, pp. 243–248). Association for Consumer Research.

Schermerhorn, P., Scheutz, M., & Crowell, C. R. (2008). Robot social presence and gender. In T. Fong, K. Dautenhahn, M. Scheutz, & Y. Demiris (Eds.), *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction - HRI '08* (pp. 263–270). ACM Press. https://doi.org/10.1145/1349822.1349857

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: a psychological approach. *Universal Access in the Information Society*, *4*(2), 146–155. https://doi.org/10.1007/s10209-005-0118-1

Severson, R. L., & Carlson, S. M. (2010). Behaving as or behaving as if? Children's conceptions of personified robots and the emergence of a new ontological category. *Neural Networks*, *23*(8-9), 1099–1103. https://doi.org/10.1016/j.neunet.2010.08.014

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, *16*(4), 337–351. https://doi.org/10.1162/pres.16.4.337

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. Springer.

Skantze, G., Hjalmarsson, A., & Oertel, C. (2014). Turn-taking, feedback and joint attention in situated human–robot interaction. *Speech Communication*, *65*, 50–66. https://doi.org/10.1016/j.specom.2014.05.005

Sosis, R. H. (1974). Internal-external control and the perception of responsibility of another for an accident. *Journal of Personality and Social Psychology*, *30*(3), 393–399. https://doi.org/10.1037/h0036893

Spence, P. R., Westerman, D., Edwards, C., & Edwards, A [Autumn] (2014). Welcoming our robot overlords: Initial expectations about interaction with a robot. *Communication Research Reports*, *31*(3), 272–280. https://doi.org/10.1080/08824096.2014.924337

Sugiyama, S. (Ed.). (2019). *Advances in human and social aspects of technology (AHSAT) book series*. *Human behavior and another kind in consciousness: Emerging research and opportunities*. IGI Global. https://doi.org/10.4018/AHSAT

Sullins, J. P. (2006). When is a robot a moral agent? *International Review of Information Ethics*, *6*, 23–30.

Sundar, S. S., & Nass, C. I. (2000). Source orientation in human-computer interaction. *Communication Research*, *27*(6), 683–703. https://doi.org/10.1177/009365000027006001

Sundar, S. S., Waddell, T. F., & Jung, E. H. (2016). The Hollywood robot syndrome: Media effects on older adults' attitudes toward robots and adoption intentions. In C. Bartneck, Y. Nagai, A. Paiva, & S. Sabanovic (Eds.), *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction - HRI '16* (pp. 343–350). IEEE. https://doi.org/10.1109/HRI.2016.7451771

Syrdal, D. S., Dautenhahn, K., Woods, S. N., Walters, M. L., & Koay, K. L. (2007). Looking good? Appearance preferences and robot personality inferences at zero acquaintance. In A. Tapus, M. Michalowski, & S. Sabanovic (Eds.), *Papers from the AAAI Spring symposium: Multidisciplinary collaboration for socially assistive robotics* (pp. 86–92). AAAI Press.

Szczuka, J. M., & Krämer, N. C. (2017). Not only the lonely - How men explicitly and implicitly evaluate the attractiveness of sex robots in comparison to the attractiveness of women, and personal characteristics influencing this evaluation. *Multimodal Technologies and Interaction*, *1*(1), Article 3. https://doi.org/10.3390/mti1010003

Tan, S. C. G., & Nareyek, A. (2009). Integrating facial, gesture, and posture emotion expression for a 3D virtual agent. In Q. Mehdi, A. Elmagraphy, D. Anderson, & E. Ch'ng (Eds.), *Proceedings of the 14th International Conference on Computer Games - CGAMES '09: AI, Animation, Mobile, Interactive Multimedia, Educational and Serious Games* (23-31). The University of Wolverhampton.

Tanaka, T., Fujikake, K., Yonekawa, T., Yamagishi, M., Inagami, M., Kinoshita, F., Aoki, H., & Kanamori, H. (2017). Driver agent for encouraging safe driving behavior for the elderly. In B. Wrede, Y. Nagai, T. Komatsu, M. Hanheide, & L. Natale (Eds.), *Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17* (pp. 71–79). ACM. https://doi.org/10.1145/3125739.3125743

van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4-5), 221–235. https://doi.org/10.1016/j.poetic.2006.05.004

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Visti, T. (2018). *World Robotic Report 2018*. International Federation of Robotic.

von der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S.-H. (2010). "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Computers in Human Behavior*, *26*(6), 1641–1650. https://doi.org/10.1016/j.chb.2010.06.012

Weiss, A., Igelsböck, J., Wurhofer, D., & Tscheligi, M. (2011). Looking forward to a "robotic society"? *International Journal of Social Robotics*, *3*(2), 111–123. https://doi.org/10.1007/s12369-010-0076-5

Wilson, W., Chun, N., & Kayatani, M. (1965). Projection, attraction, and strategy choices in intergroup competition. *Journal of Personality and Social Psychology*, *2*(3), 432–435. https://doi.org/10.1037/h0022287

Winner, L. (1977). *Autonomous technology: Technics-out-of-control as a theme in political thought*. MIT press.

Woods, S. N., Dautenhahn, K., & Kaouri, C. (2005). Is someone watching me? Consideration of social facilitation effects in human-robot interaction experiments. In *Proceedings of the 2005 International Symposium on Computational Intelligence in Robotics and Automation - CIRA '05* (pp. 53–60). IEEE. https://doi.org/10.1109/CIRA.2005.1554254

Wullenkord, R., Fraune, M. R., Eyssel, F., & Sabanovic, S. (2016). Getting in Touch: How imagined, actual, and physical contact affect evaluations of robots. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication - RO-MAN '16* (pp. 980–985). IEEE. https://doi.org/10.1109/ROMAN.2016.7745228

Young, J. E., Sung, J., Voida, A., Sharlin, E., Igarashi, T., Christensen, H. I., & Grinter, R. E. (2011). Evaluating human-robot interaction: Focusing on the holistic interaction experience. *International Journal of Social Robotics*, *3*(1), 53–67. https://doi.org/10.1007/s12369-010-0081-8

DOI: 10.17185/duepublico/74074
URN: urn:nbn:de:hbz:464-20210226-153131-4